



\*\*\*\*\*

WORLDH  
(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 11:10:36 2000; MasPar time 45.17 Seconds  
Tabular output not generated. 106.424 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (1-371) from US09376430A.pap (1 of 25)  
Perfect Score: 2788  
Sequence: 1 MGRVLVWGAAVFLGGWMA.....DVTIGGTFVMDRSYVAL 371

Scoring table: PAM 150  
Gap 11

Searched: 131253 seqs, 12956647 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-issued  
1:SA\_COMB 2:SB\_COMB 3:PCT\_COMB 4:backfiles1

Statistics: Mean 33.026; Variance 152.674; scale 0.216

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	195	7.0	230	1	US-08-595-Sequence 11, Applicati	7.47e-07
2	195	7.0	230	1	US-08-052-Sequence 11, Applicati	7.47e-07
3	195	7.0	252	1	US-08-595-Sequence 9, Applicati	7.47e-07
4	195	7.0	252	1	US-08-052-Sequence 9, Applicati	7.47e-07
5	195	7.0	347	1	US-08-595-Sequence 7, Applicati	7.47e-07
6	195	7.0	347	1	US-08-052-Sequence 7, Applicati	7.47e-07
7	195	7.0	369	1	US-08-595-Sequence 4, Applicati	7.47e-07
8	195	7.0	369	1	US-08-052-Sequence 4, Applicati	7.47e-07
9	191	6.9	536	2	US-08-456-Sequence 12, Applicati	1.57e-06
10	191	6.9	536	2	US-08-164-Sequence 12, Applicati	1.57e-06
11	188	6.7	369	2	US-08-424-Sequence 2, Applicati	2.73e-06
12	188	6.7	369	3	PCT-US94-0Sequence 69, Applicati	2.73e-06
13	178	6.4	522	2	US-08-456-Sequence 10, Applicati	1.71e-05
14	178	6.4	522	2	US-08-164-Sequence 10, Applicati	1.71e-05
15	166	6.0	379	1	US-08-456-Sequence 8, Applicati	1.51e-04
16	166	6.0	379	1	US-08-164-Sequence 8, Applicati	1.51e-04
17	166	6.0	468	2	US-08-456-Sequence 7, Applicati	1.51e-04
18	166	6.0	468	2	US-08-164-Sequence 7, Applicati	1.51e-04
19	158	5.7	258	1	US-08-336-Sequence 10, Applicati	6.37e-04
20	158	5.7	258	1	US-08-850-Sequence 5, Applicati	4.45e-04
21	132	4.7	383	1	US-08-609-Sequence 2, Applicati	6.09e-02
22	113	4.1	897	1	US-07-960-Sequence 2, Applicati	1.49e-00
23	110	3.9	551	4	5198359-2 Patent No. 5198359	2.44e+00

24	110	3.9	551	4	5449756-2	Patent No. 5449756	2.44e+00
25	107	3.8	459	4	5194375-2	Patent No. 5194375	3.97e+00
26	102	3.7	380	1	US-08-609-Sequence 4, Applicatio	8.85e+00	
27	100	3.6	788	2	US-07-728-Sequence 27, Applicati	1.22e+01	
28	100	3.6	913	1	US-08-445-Sequence 4, Applicatio	1.22e+01	
29	100	3.6	919	1	US-08-336-Sequence 2, Applicatio	1.22e+01	
30	99	3.6	982	2	US-08-673-Sequence 4, Applicatio	1.42e+01	
31	99	3.6	1073	1	US-08-241-Sequence 1, Applicatio	1.42e+01	
32	99	3.6	1073	1	US-07-695-Sequence 1, Applicatio	1.42e+01	
33	99	3.6	1091	1	US-08-241-Sequence 3, Applicatio	1.42e+01	
34	99	3.6	1091	1	US-07-695-Sequence 3, Applicatio	1.42e+01	
35	97	3.5	539	4	5198359-4	Patent No. 5198359	1.95e+01
36	97	3.5	539	4	5449756-4	Patent No. 5449756	1.95e+01
37	95	3.4	141	1	US-07-695-Sequence 5, Applicatio	2.67e+01	
38	95	3.4	149	1	US-08-241-Sequence 7, Applicatio	2.67e+01	
39	95	3.4	149	1	US-07-695-Sequence 16, Applicati	2.67e+01	
40	96	3.4	505	1	US-08-317-Sequence 16, Applicati	2.28e+01	
41	96	3.4	505	1	US-08-149-Sequence 16, Applicati	2.28e+01	
42	96	3.4	799	1	US-08-054-Sequence 2, Applicatio	2.28e+01	
43	96	3.4	822	2	US-08-474-Sequence 9, Applicatio	2.28e+01	
44	96	3.4	822	2	US-08-474-Sequence 9, Applicatio	2.28e+01	
45	96	3.4	822	2	US-08-472-Sequence 8, Applicatio	2.28e+01	

## ALIGNMENTS

RESULT 1  
ID US-08-595-974-11 STANDARD; PRT; 230 AA.  
XX  
AC  
XX  
XX  
DT  
XX  
DE Sequence 11, Application US/08595974  
XX  
CC Sequence 11, Application US/08595974  
CC Patent No. 5705608  
CC GENERAL INFORMATION:  
CC APPLICANT: SUGAMURA, KAZUO  
CC APPLICANT: TAKESHITA, TOSHIKAZU  
CC APPLICANT: ASANO, HIRONOBU  
CC APPLICANT: NAKAMURA, MASATAKA  
CC APPLICANT: SHIMAMURA, TOSHIRO  
CC APPLICANT: SUZUKI, MANABU  
CC APPLICANT: HAMURO, JUNJI  
CC TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE  
CC NUMBER OF SEQUENCES: 21  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: OBLON, SPIVAK, MCCELLELAND, MAIER & NEUSTADT,  
CC STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CC CITY: Arlington  
CC STATE: Virginia  
CC COUNTRY: U.S.A.  
CC ZIP: 22202  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA: US/08/595,974  
CC FILING DATE: 06-FEB-1996  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/052,205  
CC FILING DATE: 22-APR-1993  
CC APPLICATION NUMBER: JP 104947/1992  
CC FILING DATE: 23-APR-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Oblon, No. 5705608man F.  
CC REGISTRATION NUMBER: 24,618  
CC REFERENCE/DOCKET NUMBER: 10-615-0X



ID Q9WXR2 PRELIMINARY; PRT; 560 AA.  
 AC Q9WXR2;  
 DT 01-NOV-1999 (TRENBLrel. 12, Created)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 DE OLIGOPEPTIDE ABC TRANSPORTER, PERIPLASMIC OLIGOPEPTIDE-BINDING  
 DE PROTEIN, PUTATIVE.  
 GN TM0056.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogales; Thermotoga.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,  
 RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,  
 RA McDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,  
 RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,  
 RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,  
 RA SMITH H.O., VENTER J.C., FRASER C.M.,  
 RA "Evidence for lateral gene transfer between Archaea and bacteria from  
 RA genome sequence of Thermotoga maritima."  
 RA Nature 399:323-329(1999).  
 RP SEQUENCE FROM N.A.  
 RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,  
 RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,  
 RA McDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,  
 RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,  
 RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,  
 RA SMITH H.O., VENTER J.C., FRASER C.M.,  
 RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF001804; AAD36666.1;  
 SQ SEQUENCE 660 AA; 76186 MW; 5310F61D CRC32;

Query Match 2.2%; Score 8; DB 2; Length 660;  
 Best Local Similarity 100.0%; Pred. No. 5,006-02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 70 PEEPLVQ 77  
 QY 308 PEEPLVQ 315

RESULT 8  
 ID Q15550 PRELIMINARY; PRT; 66 AA.  
 AC Q15550;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DE ALTERNATIVELY SPLICED TAU (FRAGMENT).  
 GN TAU.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ANDREADIS A., BROWN W.M., KOSIK K.S.,  
 RA "Structure and novel exons of the human tau gene."  
 RA Biochemistry 31:10626-10633(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA ANDREADIS A.,  
 RA "The exon trapping assay partly discriminates against alternatively  
 RA spliced exons."  
 RA Nucleic Acids Res. 21:2217-2221(1993).  
 DR EMBL; AF047859; -; NOT\_ANNOTATED\_CDS.  
 KW Alternative splicing.  
 FT NON\_TER  
 SQ SEQUENCE 66 AA; 6816 MW; 56DB9D2E CRC32;

Query Match 1.9%; Score 7; DB 4; Length 66;  
 1.9%; Score 7; DB 2; Length 120;

Best Local Similarity 100.0%; Pred. No. 5,156+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 42 PSSPKHV 48  
 QY 119 PSSPKHV 125  
 RESULT 9  
 ID Q9Y4X3 PRELIMINARY; PRT; 112 AA.  
 AC Q9Y4X3;  
 DT 01-NOV-1999 (TRENBLrel. 12, Created)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 DE CCL27 CHEMOKINE.  
 GN SCY237.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ZABALLOS A., GUTIERREZ J., MARQUEZ G., HROMAS R.,  
 RA "CCL27, the human homologue of murine, ALP Chemokine."  
 RA Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ243542; CAB46983.1;  
 SQ SEQUENCE 112 AA; 12618 MW; 652C9563 CRC32;

Query Match 1.9%; Score 7; DB 4; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 5,186+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 9 SLLLSL 15  
 QY 245 SLLLSL 251

RESULT 10  
 ID Q9X1T5 PRELIMINARY; PRT; 120 AA.  
 AC Q9X1T5;  
 DT 01-NOV-1999 (TRENBLrel. 12, Created)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 DE HYPOTHETICAL 14.0 KD PROTEIN.  
 GN TM1599.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogales; Thermotoga.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,  
 RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,  
 RA McDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,  
 RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,  
 RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,  
 RA SMITH H.O., VENTER J.C., FRASER C.M.,  
 RA "Evidence for lateral gene transfer between Archaea and bacteria from  
 RA genome sequence of Thermotoga maritima."  
 RA Nature 399:323-329(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,  
 RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,  
 RA McDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,  
 RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,  
 RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,  
 RA SMITH H.O., VENTER J.C., FRASER C.M.,  
 RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF001804; AAD36666.1;  
 SQ SEQUENCE 120 AA; 14025 MW; AC553C66 CRC32;

Query Match 1.9%; Score 7; DB 2; Length 120;

US-09-376-430-2.rspt

Thu May 11 06:50:39 2000

Best Local Similarity 100.0%; Pred. No. 5.16e+00; Mismatches 0; Conservative 0; Indels 0; Gaps 0;

Db 75 VSLLLS 81  
QY 244 VSLLLS 250  
RESULT 11  
ID O60542 PRELIMINARY; PRT; 156 AA.  
AC O60542;  
DT 01-AUG-1998 (TRENBLrel. 07, Created)  
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)  
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
DE PERSEPHIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 98150950.  
RA MILBRANDT J., DE SAUVAGE F.J., FAHRNER T.J., BALOH R.H., LEITNER M.L.,  
RA TANSBY M.G., LAMPE P.A., HEUCKERTH R.O., KOTZBAUER P.T.,  
RA STUBERGER K.S., GOLDEN J.P., DAVIES J.A., VEJSADA R., KATO A.C.,  
RA HYNES M., SHERMAN D., NISHIMURA M., WANG L.-C., VANDLEN R., MOFFAT B.,  
RA KLEIN R.D., POULSEN K., GRAY C., GARCES A., HENDERSON C.E.,  
RA PHILLIPS H.S., JOHNSON E.M.;  
RT "Persephin, a novel neurotrophic factor related to GDNF and  
RT neurturin";  
RL Neuron 20:245-253(1998).  
DR EMBL; AF040962; AAC39640.1; -;  
DR HSP; Q07731; IAGO.  
SQ SEQUENCE 156 AA; 16600 MW; 80D8A1E4 CRC32;

Query Match 1.9%; Score 7; DB 4; Length 156;  
Best Local Similarity 100.0%; Pred. No. 5.16e+00;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 10 SLLLSL 16  
QY 245 SLLLSL 251  
RESULT 12  
ID Q925W2 PRELIMINARY; PRT; 176 AA.  
AC Q925W2;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE ORPHO-HALOGENASE 1,2-DIOXYGENASE BETA-ISP PROTEIN OHBA.  
GN OHBA.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;  
OC Pseudomonas.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-142;  
RA TSOI T.V., PLOTNIKOVA E.G., COLE J.R., GUERIN W.F., BAGDASARIAN M.,  
RA TIEDJE J.M.;  
RT "Cloning, expression and nucleotide sequence of the Pseudomonas  
RT aeruginosa strain 142 ohb genes for oxygenolytic ortho-dehalogenation  
RT of halobenzenes";  
RL Appl. Environ. Microbiol. 65:0-0(1999).  
DR EMBL; AF121970; AAD20005.1; -;  
KW DIOXYGENASE.  
SQ SEQUENCE 176 AA; 20252 MW; EBD8F11 CRC32;

Query Match 1.9%; Score 7; DB 2; Length 176;  
Best Local Similarity 100.0%; Pred. No. 5.16e+00;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 108 VRYKAME 114  
QY 245 VRYKAME 114

QY 185 VRYKAME 191  
RESULT 13  
ID Q38901 PRELIMINARY; PRT; 176 AA.  
AC Q38901;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
DE CYTOSOLIC CYCLOPHILIN (EC 5.2.1.8).  
GN ROC2.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-COLUMBIA;  
RX MEDLINE; 98088013.  
RA CHOU T.T., GASSER C.S.;  
RT "Characterization of the cyclophilin gene family of Arabidopsis  
RT thaliana and phylogenetic analysis of known cyclophilin proteins";  
RL Plant Mol. Biol. 35:873-892(1997).  
CC -1- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC  
CC PEPTIDE BONDS IN OLIGOPEPTIDES.  
DR EMBL; U04000; AAB96833.1; -;  
DR HSP; P05092; 3CYS.  
DR MENDEL; 6283; Arath; 1040; 6283.  
DR PROSITE; PS00170; CSA\_PPIASE\_1; 1.  
DR PFAM; PF00160; pro\_isomerase; 1.  
DR PRINTS; PR00153; CSAPPISMASE.  
KW Isomerase; Rotamase.  
SQ SEQUENCE 176 AA; 18906 MW; 07BB7AB3 CRC32;

Query Match 1.9%; Score 7; DB 10; Length 176;  
Best Local Similarity 100.0%; Pred. No. 5.16e+00;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 146 VKAMEDV 152  
QY 187 VKAMEDV 193  
RESULT 14  
ID Q45498 PRELIMINARY; PRT; 212 AA.  
AC Q45498;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE HYPOTHETICAL 24.6 KD PROTEIN.  
GN YKTB.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/staphylococcus group; Bacillus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-1168;  
RX MEDLINE; 90368558.  
RA HEMILA H., PALVA A., PAULIN L., ARVIDSON S., PALVA I.;  
RT "secretory S complex of Bacillus subtilis: sequence analysis and  
RT identity to pyruvate dehydrogenase";  
RL J. Bacteriol. 172:5052-5063(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-1168;  
RX MEDLINE; 97124187.  
RA WINTERS P., CALDWELL R., ENFIELD L., FERRARI E.;  
RT "The ampS-*np*E (124 degrees-127 degrees) region of the Bacillus  
RT subtilis 168 chromosome: sequencing of a 27 kb segment and  
RT identification of several genes in the area";  
RL Microbiology 142:0-0(0).  
RN [3]

CC ADDRESS: P.C.  
CC STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CC CITY: Arlington  
CC STATE: Virginia  
CC COUNTRY: U.S.A.  
CC ZIP: 22202

CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent in Release #1.0, Version #1.25  
CC CURRENT APPLICATION NUMBER: US/08/595,974  
CC FILING DATE: 06 FEB 1996  
CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/052,205  
CC FILING DATE: 22-APR-1993  
CC APPLICATION NUMBER: JP 104947/1992  
CC FILING DATE: 23-APR-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: OBLON, NO. 5705608man F.  
CC REGISTRATION NUMBER: 24,618  
CC REFERENCE/DOCKET NUMBER: 10-615-0X  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (703) 413-3000  
CC TELEFAX: (703) 413-2220

CC INFORMATION FOR SEQ ID NO: 7:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 347 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 347 AA; 39920 MW; 666553 CN;

Query Match 7.0%; Score 195; DB 1; Length 347;  
Best Local Similarity 25.9%; Pred. No. 7.47e-07;  
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;

Db 37 EVQCFVFNVEYMNCTNSSEPPQTNLTLYHYKNSDNDKVKQSHYLFSEITSGCOLQ 96  
QY 31 QIQIIFNLEVTQVWNAKYSR-TNLTTHYRF-NGD-EAYDQCTNVLQEGHTSGCLLD 87  
Db 97 KKEIHYQTVFVVLQDPPRPRQATQMLKQNLVWPAPENLTLYHYKNSDNDKVKQSHYLFSEITSGCOLQ 96  
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VIYL-KPSSPKHVRP-SWHQDAVTVTCSDL 142  
Db 157 LNHCLHLYQVYRTDWDHSHWTEQSDVYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 215  
QY 143 YGD-LLYEVOYRSPFDTEW-QSKQENTCNVTIEGLDAEKCYSFWRVVKAMEDYVGPDTYP 200  
Db 216 SEWSPHIW 224  
QY 201 SDWSEVTCW 209

RESULT 5  
ID US-08-052-205-7 STANDARD; PRT; 347 AA.  
XX xxxxxx  
XX  
XX  
XX

Sequence 7, Application US/08052205  
Sequence 7, Application US/08052205  
Patent No. 510259  
GENERAL INFORMATION:  
CC APPLICANT: SUGAMURA, KAZUO  
CC APPLICANT: TAKESHITA, TOSHIKAZU  
CC APPLICANT: ASAO, HIROSHI  
CC APPLICANT: NAKAMURA, MASATAKA

CC APPLICANT: SHIMAMURA, TOSHIRO  
CC APPLICANT: SUZUKI, MANABU  
CC APPLICANT: HAMURO, JUNJI  
CC TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE  
CC NUMBER OF SEQUENCES: 21  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
CC ADDRESSEE: P.C.  
CC STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CC CITY: Arlington  
CC STATE: Virginia  
CC COUNTRY: U.S.A.  
CC ZIP: 22202

CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent in Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/052,205  
CC FILING DATE: 19930422  
CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: JP 104947/1992  
CC FILING DATE: 23-APR-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: OBLON, NO. 5510259man F.  
CC REGISTRATION NUMBER: 24,618  
CC REFERENCE/DOCKET NUMBER: 10-615-0X  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (703) 413-3000  
CC TELEFAX: (703) 413-2220

CC INFORMATION FOR SEQ ID NO: 7:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 347 amino acids  
CC TYPE: AMINO ACID  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 347 AA; 39920 MW; 666553 CN;

Query Match 7.0%; Score 195; DB 1; Length 347;  
Best Local Similarity 25.9%; Pred. No. 7.47e-07;  
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;

Db 37 EVQCFVFNVEYMNCTNSSEPPQTNLTLYHYKNSDNDKVKQSHYLFSEITSGCOLQ 96  
QY 31 QIQIIFNLEVTQVWNAKYSR-TNLTTHYRF-NGD-EAYDQCTNVLQEGHTSGCLLD 87  
Db 97 KKEIHYQTVFVVLQDPPRPRQATQMLKQNLVWPAPENLTLYHYKNSDNDKVKQSHYLFSEITSGCOLQ 96  
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VIYL-KPSSPKHVRP-SWHQDAVTVTCSDL 142  
Db 157 LNHCLHLYQVYRTDWDHSHWTEQSDVYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 215  
QY 143 YGD-LLYEVOYRSPFDTEW-QSKQENTCNVTIEGLDAEKCYSFWRVVKAMEDYVGPDTYP 200  
Db 216 SEWSPHIW 224  
QY 201 SDWSEVTCW 209

RESULT 7  
ID US-08-595-974-4 STANDARD; PRT; 369 AA.  
XX xxxxxx  
XX  
XX  
XX

Sequence 4, Application US/08595974  
Sequence 4, Application US/08595974  
Patent No. 5705608

```

XX sequence 9, Application US/08052205
DE
XX
XX
XX sequence 9, Application US/08052205
XX Patent No. 5510259
XX
XX GENERAL INFORMATION:
XX
XX APPLICANT: SUGAMURA, KAZUO
XX
XX APPLICANT: TAKESHITA, TOSHIKAZU
XX
XX APPLICANT: ASAO, HIRONOBU
XX
XX APPLICANT: NAKAMURA, MASATAKA
XX
XX APPLICANT: SHIMAMURA, TOSHIRO
XX
XX APPLICANT: SUZUKI, MANABU
XX
XX APPLICANT: HAMURO, JUNJI
XX
XX TITLE OF INVENTION: HUMAN IL-2
XX
XX NUMBER OF SEQUENCES: 21
XX
XX CORRESPONDENCE ADDRESSES:
XX
XX ADDRESSEE: OBLON, SPIVAK, MCCOY
XX
XX ADDRESSEE: P.C.
XX
XX STREET: 1755 S. Jefferson
XX
XX CITY: Arlington
XX
XX STATE: Virginia
XX
XX COUNTRY: U.S.A.
XX

```

RESULT 5  
ID US-08-595-974-7 STANDARD: PRT: 347 AA.

Sequence 7, Application US/08595974  
Sequence 7, Application US/08595974  
Patent No. 5705608  
GENERAL INFORMATION:  
APPLICANT: SUGAMURA, KAZUO  
APPLICANT: TAKESHITA, TOSHIKAZU  
APPLICANT: ASAO, HIRONOBU  
APPLICANT: NAKAMURA, MASATAKA  
APPLICANT: SHIMAMURA, TOSHIRO  
APPLICANT: SUZUKI, MANABU  
APPLICANT: HAMURO, JUNJI  
TITLE OF INVENTION: HUMAN IL-2  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCOY



GENERAL INFORMATION:  
 APPLICANT: SUGAMURA, KAZUO  
 APPLICANT: TAKESHITA, TOSHIKAZU  
 APPLICANT: ASAO, HIRONOBU  
 APPLICANT: NAKAMURA, MASATAKA  
 APPLICANT: SHIMAMURA, TOSHIRO  
 APPLICANT: SUZUKI, MANABU  
 APPLICANT: HAMURO, JUNJI  
 TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
 STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: U.S.A.  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/595,974  
 FILING DATE: 06-FEB-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/052,205  
 FILING DATE: 22-APR-1993  
 APPLICATION NUMBER: JP 104947/1992  
 FILING DATE: 23-APR-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Oblon, No. 5705608man F.  
 REGISTRATION NUMBER: 24,618  
 REFERENCE/DOCKET NUMBER: 10-615-0X  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 413-3000  
 TELEFAX: (703) 413-2220  
 TELEX: 248855 OPAT UR  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 369 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE 369 AA; 42287 MW; 754355 CN;

Sequence 4, Application US/08052205  
Sequence 4, Application US/08052205  
Patent No. 5510259  
GENERAL INFORMATION:  
APPLICANT: SUGAMURA, KAZUO  
APPLICANT: TAKESHITA, TOSHIKAZU  
APPLICANT: ASAO, HIRONOBU  
APPLICANT: NAKAMURA, MASATAKA  
APPLICANT: SHIMAMURA, TOSHIRO  
APPLICANT: SUZUKI, MANABU  
APPLICANT: HAMURO, JUNJI  
TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADRESSEE: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/052,205  
FILING DATE: 19930422  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 104947/1992  
FILING DATE: 23-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5510259man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 10-615-0X  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 369 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE 369 AA; 42287 MW; 754355 CN;

```

Query Match      7.0%  Score 195;  DB 1;  Length 369;
Best Local Similarity 25.9%;  Pred. No. 7.47e-07;
Matches 49;  Conservative 52;  Mismatches 77;  Indels 11;  Gaps 11;

Db 59 EVQCFEVEVVEYMCNTWNSSEPOPTNLTHRYWYKNSDNDKVQKSHYLFSEETISGCQLQ 118
    : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 31 QIQIIYFNLETVQVWTWASYSR-TNLTHRYRF-NGD-EAYDQCTNYLLQEGHTSGCLLD 87

Db 119 KKEITHLYCTFVQLODPRPRRATOMLQNLQVLPWAPENLTLKSESOLELWNNRF 178
    : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYVL-KPSSPKHVRF-SWHQDAVTVTCSDL 142

Db 179 LNHCEHLVQVYRTDMSHWTEQSVYDRHKFSPVSDGQKRYTFYRVSF-FNPLGCSAQHW 237
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 143 YGD-LLYEVQYRSPFDTEW-QSKQENTCNVTIEGLDAEKCSFWVRVKAMEDVYGPDTYP 200

Db 238 SEMSHPIHW 246
    | : | | | |
Qy 201 SDWSVTCW 209

RESULT 8
ID US-08-052-205-4 STANDARD; PRt; 369 AA.
XX
AC xxxxxx

```

RP SEQUENCE FROM N.A.  
 RC STRAIN=1168;  
 RX MEDLINE: 97144523.  
 RA HENRIQUES A.O., BRYAN E.M., BEALL B.W., MORAN C.P. JR.:  
 RT "cse15, cse60, and cse22 are new members of mother-cell-specific  
 RL sporulation regulons in *Bacillus subtilis*."  
 RN J. Bacteriol. 179:389-398(1997).  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE: 98044033.  
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
 RA AZEVEDO V., BERTELO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,  
 RA BOURISS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,  
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,  
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,  
 RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMERSON P.T.,  
 RA ENRIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,  
 RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,  
 RA GHIM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,  
 RA GUSEPPI G., GUY B.J., HAGA K., HAIECH J., HARWOOD C.R., HENAUT A.,  
 RA HILBERT H., HOLZAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,  
 RA JORIS B., KARAMATA D., KASAHARA Y., KLAER-BLANCHARD M., KLEIN C.,  
 RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M.,  
 RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,  
 RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,  
 RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,  
 RA NOONE D., O'REILLY M., OGAWA K., OGIWARA A., OUDEGA B., PARK S.H.,  
 RA PARRO V., POHL T.M., PORTETELLE D., FORMOLIK S., PRESCOTT A.M.,  
 RA PRESECAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,  
 RA RIEGER M., RIVOLTA C., ROCHE E., ROCHE B., ROSE M., SADAIE Y.,  
 RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOFFONE F.,  
 RA SEKIGUCHI J., SEKOWSKA A., SEROR S.J., SERROR P., SHIN B.S., SOLDI B.,  
 RA SOROKIN A., TACCONE E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,  
 RA TAKEUCHI M., TAKAKOSHI A., TANAKA T., TERPSTRA P., TOGNONI A.,  
 RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,  
 RA VIARI A., WAMBUTT R., WEDLER E., WEDLER H., WEITZENEGER T.,  
 RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUNOTO K., YATA K.,  
 RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.,  
 RT The complete genome sequence of the gram-positive bacterium *Bacillus*  
 RT *subtilis*.  
 RL Nature 390:249-256(1997).  
 DR EMBL: AF012285; AAC24939.1; -  
 DR EMBL: 299111; CAB13338.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 212 AA; 24616 MW; 824502E2 CRC32;

Query Match 1.9%; Score 7; DB 2; Length 212;  
 Best Local Similarity 100.0%; Pred. No. 5.16e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 15 TIEGLDA 21  
 QY 171 TIEGLDA 177

RESULT 15  
 ID 080097 PRELIMINARY; PRT; 216 AA.  
 AC 080097;  
 DT 01-NOV-1998 (TIEGLDA 08, Created)  
 DT 01-NOV-1998 (TIEGLDA 08, Last sequence update)  
 DT 01-NOV-1998 (TIEGLDA 08, Last annotation update)  
 DE ORF 59  
 OS bacteriophage phi PVL  
 OC Viruses.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 98067870.  
 RA KANEKO J., KIMURA T., KAWAKAMI Y., TOMITA T., KAMIO Y.,  
 ET "Fancin" (a leucocidin gene) in a phage-like particle isolated  
 ET from mitomycin C-treated *Staphylococcus aureus* V8 (ATCC 49775)."  
 RL Bacteriophage J. Biochem. 61:1960-1962(1997).  
 DR EMBL: AB009866; BAA31932.1; -

SQ SEQUENCE 216 AA; 26026 MW; A10682DF CRC32;  
 Query Match 1.9%; Score 7; DB 9; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 5.16e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 105 RVRAMED 111  
 QY 186 RVRAMED 192  
 Search completed: Wed May 10 14:38:35 2000  
 Job time : 267 secs.







QY 141 LSYGDLLEYOYRSPFDTW-QSKQENTCNVTIEGLDAEKCYSFWRVWKAMEDVYGPDTY 199  
Db 238 WSKSQPWH 247  
QY 200 PSDNSEVTCW 209  
RESULT 13  
ID US-08-456-489B-10 STANDARD; PRT; 522 AA.  
XX xxxxxx  
AC  
XX  
XX  
XX  
XX  
DE Sequence 10, Application US/08456489B  
XX  
CC Sequence 10, Application US/08456489B  
CC Patent No. 5962269  
CC GENERAL INFORMATION:  
CC APPLICANT: Renauld, Jean-Christophe; Druetz, Catherine; Van Snick,  
CC APPLICANT: Jacques  
CC TITLE OF INVENTION: Nucleic Acid Sequences Coding For Or  
CC TITLE OF INVENTION: Complementary To Nucleic Acid Sequences Coding For Interle  
CC NUMBER OF SEQUENCES: 12  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Fulbright & Jaworski L.L.P.  
CC STREET: 666 Fifth Avenue  
CC CITY: New York City  
CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10103  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
CC COMPUTER: IBM PS/2  
CC OPERATING SYSTEM: PC-DOS  
CC SOFTWARE: Wordperfect  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/456,489B  
CC FILING DATE: 1-JUNE-1995  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 08/164,614  
CC FILING DATE: 8-DECEMBER-1993  
CC PRIOR APPLICATION NUMBER: 07/847,347  
CC FILING DATE: 09-MARCH-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Schofield, Mary Anne  
CC REGISTRATION NUMBER: 36,669  
CC REFERENCE/DOCKET NUMBER: LUD 5264.2  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 318-3000  
CC TELEFAX: (212) 752-5958  
CC INFORMATION FOR SEQ ID NO: 10:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 522 amino acid residues  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC SEQUENCE 522 AA; 57319 MW; 1432274 CN;  
Query Match 6.4%; Score 178; DB 2; Length 522;  
Best Local Similarity 27.2%; Pred. No. 1.71e-05;  
Matches 62; Conservative 54; Mismatches 87; Indels 25; Gaps 21;  
Db 178 LLSYELAFKKQ-EEAWEQAQRHDHVGVTWILLEAFVDPGFTHEARLVOMATLEDDVV 236  
QY 146 LL-YYOYRSPFDTW-QSKQ-ENTCNVT--I-EGDLAEKCYSF--WVRV-KA-MED-VY 194  
Db 237 EERYTCGQNSE---NSQPCVCFQAPQROGRLPPMGWPGN-TLVAVSIFLLTGPYTL-LF 291  
QY 195 GPDYPSDNSEVTCWQGE-IRDAEAETP-TPPKPKLSKFLISSAILLMVSLLSLW 252  
Db 292 KLSPRVKRIFYQNVSPAMFFQPLYSVHNGNFQTMGAHRAGVLLSQDCAGTPOGALEPC 351

QY 253 KLM-RVKFLIPSPDPKSIFFGLFEIHGQNFQEWITTONVAHL-HKMAQAEQES-GPE 309  
Db 352 VQETALLTCGPAPKMSKVALEEEQEPG-T-RLFGN-LSSEEDVLPAG 396  
QY 310 EPLVVQLAKTEASPRMLDPQTEKEASGSLQPLPQLOGGDVVTIG 357  
RESULT 14  
ID US-08-164-614A-10 STANDARD; PRT; 522 AA.  
XX xxxxxx  
AC  
XX  
XX  
XX  
XX  
DE Sequence 10, Application US/08164614A  
XX  
CC Sequence 10, Application US/08164614A  
CC Patent No. 5789237  
CC GENERAL INFORMATION:  
CC APPLICANT: Renauld, Jean-Christophe  
CC APPLICANT: Druetz, Catherine  
CC APPLICANT: Van Snick, Jacques  
CC TITLE OF INVENTION: Nucleic Acid Sequences Coding For  
CC TITLE OF INVENTION: Or Complementary To Nucleic Acid Sequences Coding For  
CC TITLE OF INVENTION: Interleukin 9 Receptor  
CC NUMBER OF SEQUENCES: 12  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Felfe & Lynch  
CC STREET: 805 Third Avenue  
CC CITY: New York City  
CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10022  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
CC COMPUTER: IBM PS/2  
CC OPERATING SYSTEM: PC-DOS  
CC SOFTWARE: Wordperfect  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/164,614A  
CC FILING DATE: 8-DECEMBER-1993  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 07/847,347  
CC FILING DATE: 09-MARCH-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Hanson, No. 5789237man D.  
CC REGISTRATION NUMBER: 30,946  
CC REFERENCE/DOCKET NUMBER: LUD 5264.1  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 688-9200  
CC TELEFAX: (212) 838-3884  
CC INFORMATION FOR SEQ ID NO: 10:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 522 amino acid residues  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC SEQUENCE 522 AA; 57319 MW; 1432274 CN;  
Query Match 6.4%; Score 178; DB 1; Length 522;  
Best Local Similarity 27.2%; Pred. No. 1.71e-05;  
Matches 62; Conservative 54; Mismatches 87; Indels 25; Gaps 21;  
Db 178 LLSYELAFKKQ-EEAWEQAQRHDHVGVTWILLEAFVDPGFTHEARLVOMATLEDDVV 236  
QY 146 LL-YYOYRSPFDTW-QSKQ-ENTCNVT--I-EGDLAEKCYSF--WVRV-KA-MED-VY 194  
Db 237 EERYTCGQNSE---NSQPCVCFQAPQROGRLPPMGWPGN-TLVAVSIFLLTGPYTL-LF 291  
QY 195 GPDYPSDNSEVTCWQGE-IRDAEAETP-TPPKPKLSKFLISSAILLMVSLLSLW 252  
Db 292 KLSPRVKRIFYQNVSPAMFFQPLYSVHNGNFQTMGAHRAGVLLSQDCAGTPOGALEPC 351

QY 253 KLV-RVKFLIPSPDPKIFPGLEIFGHNQFQWITDQVAHL-HKMGAEQES-GPE 309  
 DB 352 VQETALLTCGPAPKPKNSVALSEEGPG-T-RLPGN-LSEEDVLPAG 396  
 QY 310 EPLVVLAKTEAPRMDPQTEKEASGSLQLPHQPLQGGDWTIG 357

RESULT 15  
 ID US-08-456-489B-8 STANDARD; PRT: 379 AA.

AC XXXXX

DE Sequence 8, Application US/08456489B

CC Sequence 8, Application US/08456489B

CC Patent No. 5962269

CC GENERAL INFORMATION:

CC APPLICANT: Renault, Jean-Christophe; Druetz, Catherine; Van Snick,

CC APPLICANT: Jacques

CC TITLE OF INVENTION: Nucleic Acid Sequences Coding For Or

CC TITLE OF INVENTION: Complementary To Nucleic Acid Sequences Coding For Interle

CC NUMBER OF SEQUENCES: 12

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Fulbright & Jaworski L.L.P.

CC STREET: 666 Fifth Avenue

CC CITY: New York City

CC STATE: New York

CC COUNTRY: USA

CC ZIP: 10103

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

CC COMPUTER: IBM PS/2

CC OPERATING SYSTEM: PC-DOS

CC SOFTWARE: Wordperfect

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/456,489B

CC FILING DATE: 1-JUNE-1995

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: 08/164,614

CC FILING DATE: 8-DECEMBER-1993

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: 07/847,347

CC FILING DATE: 09-MARCH-1992

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Schofield, Mary Anne

CC REGISTRATION NUMBER: 36,669

CC REFERENCE/DOCKET NUMBER: LUD 5264.2

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (212) 318-3000

CC TELEFAX: (212) 752-5958

CC INFORMATION FOR SEQ ID NO: 8:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 379 amino acid residues

CC TYPE: amino acid

CC TOPOLOGY: linear

CC SEQUENCE 379 AA: 43053 MW: 791283 CN;

Query Match

Best Local Similarity 24.7%; Score 166; DB 2; Length 379;

Matches 20; Conservative 28; Mismatches 31; Indels 2; Gaps 2;

DB 274 ILVVVFIFLLTGTVHL-LFKLSRLKRIYQNPSPFAFFHPLYSVYHGDQSWTGARR 332

QY 233 ILISSAILMLVSLLLSLKLN-RVKFLIPSPDPKIFPGLEIFGHNQFQWITDQ 291

DB 333 AGPQARQNGVSTSSAGSESSI 353

QY 292 NVAHLHKMGAEQESGPEPL 312

Search completed: Wed May 10 11:11:29 2000  
 Job time : 53 secs.

1. *Pharmaceutical industry* – The pharmaceutical industry is the largest of the three industries, with sales of \$10.5 billion in 1997. It is the only industry that has not experienced a decline in sales since 1990. The industry is dominated by a few large firms, with the top five firms accounting for 40% of sales. The industry is also characterized by high R&D spending, with the top five firms accounting for 60% of total R&D spending.

\*\*\*\*\*  
 M P S R C H  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 11:03:46 2000; MasPar time 82.75 Seconds  
 Tabular output not generated. 136.543 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (1-371) from US09376430A.ppt (1 of 25)  
 Perfect Score: 2788  
 Sequence: 1 MGRLLVLLGCAAVFLGGWMA.....DVTGGTFVANDRSYVAL 371

Scoring table: PAM 150  
 Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: swiss-prot38  
 1:swissprot

Statistics: Mean 49.107; Variance 86.776; scale 0.566

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	219	7.9	379	1	CYRG_BOVIN CYTOKINE RECEPTOR COMM	3.51e-21
2	210	7.5	373	1	CYRG_CANEA CYTOKINE RECEPTOR COMM	1.88e-19
3	195	7.0	369	1	CYRG_HUMAN CYTOKINE RECEPTOR COMM	1.28e-16
4	188	6.7	369	1	CYRG_MOUSE CYTOKINE RECEPTOR COMM	2.56e-15
5	176	6.3	522	1	IL3R_HUMAN INTERLEUKIN-9 RECEPTOR	3.98e-13
6	166	6.0	468	1	IL3R_MOUSE INTERLEUKIN-9 RECEPTOR	2.44e-11
7	164	5.9	507	1	EPOR_RAT ERYTHROPOIETIN RECEPTOR	5.49e-11
8	163	5.8	507	1	EPOR_MOUSE ERYTHROPOIETIN RECEPTOR	8.22e-11
9	160	5.7	508	1	EPOR_HUMAN ERYTHROPOIETIN RECEPTOR	2.75e-10
10	126	4.5	490	1	CP23_MOUSE CYTOCHROME P450 2C38	1.16e-04
11	119	4.3	424	1	IL31_MOUSE CYTOCHROME P450 2C39	1.37e-03
12	113	4.1	490	1	CP24_MOUSE CYTOCHROME P450 2C39	1.05e-02
13	115	4.1	878	1	IL3B_MOUSE INTERLEUKIN-3 RECEPTOR	5.37e-03
14	113	4.1	897	1	CYC7_RAT CYTOKINE RECEPTOR COMM	1.05e-02
15	108	3.9	490	1	IL2B_HUMAN INTERLEUKIN-2 RECEPTOR	5.47e-02
16	110	3.9	551	1	IL2B_MOUSE GROWTH HORMONE RECEPTOR	2.85e-02
17	110	3.9	634	1	GHR_SHEEP GROWTH HORMONE RECEPTOR	3.95e-02
18	109	3.9	634	1	GHR_BOVIN GROWTH HORMONE RECEPTOR	2.85e-02
19	110	3.9	638	1	GHR_PIG GROWTH HORMONE RECEPTOR	2.85e-02
20	110	3.9	896	1	CYRG_MOUSE CYTOKINE RECEPTOR COMM	5.47e-02
21	108	3.9	896	1	CYRG_MOUSE C-C CHEMOKINE RECEPTOR	1.43e-01
22	105	3.8	359	1	LEUK_RAT LEUKOSIALIN PRECURSOR	1.43e-01
23	105	3.8	378	1	LEUK_RAT LEUKOSIALIN PRECURSOR	1.43e-01

24	107	3.8	427	1	IL31_HUMAN INTERLEUKIN-13 RECEPTO	7.56e-02
25	107	3.8	459	1	IL7R_HUMAN INTERLEUKIN-7 RECEPTO	7.56e-02
26	106	3.8	1073	1	HSEB_HUMAN HEAT-STABLE ENTEROTOXIN	1.04e-01
27	102	3.7	380	1	IL32_HUMAN INTERLEUKIN-13 RECEPTO	3.68e-01
28	102	3.7	386	1	LMP1_EBV LATENT MEMBRANE PROTEIN	3.68e-01
29	104	3.7	1151	1	ITAI_HUMAN INTEGRIN ALPHA-1 (LAMI	1.97e-01
30	104	3.7	1180	1	ITAI_RAT INTEGRIN ALPHA-1 PRECU	1.97e-01
31	101	3.6	184	1	RL5_THOMA 50S RIBOSOMAL PROTEIN	5.02e-01
32	101	3.6	359	1	CRK3_MOUSE PROBABLE C-C CHEMOKINE	5.02e-01
33	99	3.6	386	1	LMP1_EBV LATENT MEMBRANE PROTEIN	9.26e-01
34	101	3.6	404	1	LMP1_EBVC LATENT MEMBRANE PROTEIN	5.02e-01
35	100	3.6	669	1	YH06_HAEIN PROBABLE TRANSPORT PRO	6.83e-01
36	100	3.6	788	1	ITB6_HUMAN INTEGRIN BETA-6 PRECUR	6.83e-01
37	101	3.6	864	1	LDLR_MOUSE LOW-DENSITY LIPOPROTEIN	5.02e-01
38	100	3.6	910	1	EDD1_MOUSE EPITHELIAL DISCOIDIN D	6.83e-01
39	100	3.6	911	1	EDD1_MOUSE EPITHELIAL DISCOIDIN D	6.83e-01
40	100	3.6	913	1	EDD1_MOUSE EPITHELIAL DISCOIDIN D	6.83e-01
41	99	3.6	1073	1	ITAG_HUMAN INTEGRIN ALPHA-6 PRECU	9.26e-01
42	98	3.5	490	1	CPCC_RAT CYTOCHROME P450 2C14	1.25e-00
43	98	3.5	537	1	IL2B_RAT INTERLEUKIN-2 RECEPTOR	1.25e-00
44	97	3.5	539	1	IL2B_MOUSE INTERLEUKIN-2 RECEPTOR	1.89e+00
45	97	3.5	638	1	GHR_RABIT GROWTH HORMONE RECEPTO	1.89e+00

ALIGNMENTS

RESULT 1  
 ID CYRG\_BOVIN STANDARD: PRT: 379 AA.  
 AC Q95118;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C)  
 DE (INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64).  
 GN IL2RG.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;  
 OC Bovinae; Bos.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 96268473.  
 RA Yoo J., Stone R.T., Solinas-Toldo S., Fries R., Beattie C.W.;  
 RT "Cloning and chromosomal mapping of bovine interleukin-2 receptor  
 gamma gene".  
 RL DNA Cell Biol. 15:453-459(1996).  
 CC -!- FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF  
 CC INTERLEUKINS.  
 CC -!- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND  
 CC PROBABLY ALSO THE IL-13 RECEPTORS.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: U33748; AAB07812.1;  
 CC HSSP: P31785; ILIN;  
 CC PROSITE: PS00241; RECEPTOR\_CYTOKINES\_1;  
 CC PROSITE: PS00340; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.  
 CC PFAM: PF00041; fn3; 1  
 CC Receptor; Transmembrane; Glycoprotein; Signal.  
 CC SIGNAL 23 379  
 CC CHAIN 23 379  
 CC DOMAIN 23 269  
 CC TRANSMEM 270 290  
 CC DOMAIN 291 379  
 CC CYTOPLASMIC (POTENTIAL).  
 CC



RN [3] SEQUENCE FROM N.A., AND VARIANTS ASP-114 AND ASN-153.  
 RX MEDLINE; 94004847.  
 RA Puck J.M., Deschenes S.M., Porter J.C., Dutra A.S., Brown C.J.,  
 RA Willard H., Henthorn P.S.;  
 RT "The interleukin-2 receptor gamma chain maps to Xq13.1 and is mutated  
 RT in X-linked severe combined immunodeficiency, SCIDX1.";  
 RL Hum. Mol. Genet. 2:1099-1104(1993).  
 RN [4]  
 RN IDENTIFICATION AS A IL-4R SUBUNIT.  
 RX MEDLINE; 94090315.  
 RA Kondo M., Takeshita T., Ishii N., Nakamura M., Watanabe S.,  
 RA Arai K.-I., Sugamura K.;  
 RT "Sharing of the interleukin-2 (IL-2) receptor gamma chain between  
 RT receptors for IL-2 and IL-4.";  
 RL Science 262:1874-1877(1993).  
 RN [5]  
 RN IDENTIFICATION AS A IL-4R SUBUNIT.  
 RX MEDLINE; 94090317.  
 RA Russell S.M., Kegan A.D., Harada N., Nakamura Y., Noguchi M.,  
 RA Letand P., Friedmann M.C., Miyajima A., Puri R.K., Paul W.E.,  
 RA Leonard W.J.;  
 RT "Interleukin-2 receptor gamma chain: a functional component of the  
 RT interleukin-4 receptor.";  
 RL Science 262:1880-1883(1993).  
 RN [6]  
 RN IDENTIFICATION AS A IL-7R SUBUNIT.  
 RX MEDLINE; 94090316.  
 RA Noguchi M., Nakamura Y., Russell S.M., Ziegler S.F., Tsang M., Cao X.,  
 RA Leonard W.J.;  
 RT "Interleukin-2 receptor gamma chain: a functional component of the  
 RT interleukin-7 receptor.";  
 RL Science 262:1877-1880(1993).  
 RN [7]  
 RN 3D-STRUCTURE MODELING OF 57-248.  
 RX MEDLINE; 95111955.  
 RA Bamorough P., Hedgecock C.J., Richards W.G.;  
 RT "The interleukin-2 and interleukin-4 receptors studied by molecular  
 RT modelling.";  
 RL Structure 2:839-851(1994).  
 RN [8]  
 RN VARIANTS XSCID PHE-115; CYS-240 AND ILE-241.  
 RX MEDLINE; 94130970.  
 RA Disanto J.P., Dautry-Varsat A., Certain S., Fischer A.,  
 RA de Saint Basile G.;  
 RT "Interleukin-2 (IL-2) receptor gamma chain mutations in X-linked  
 RT severe combined immunodeficiency disease result in the loss of  
 RT high-affinity IL-2 receptor binding.";  
 RL Eur. J. Immunol. 24:475-479(1994).  
 RN [9]  
 RN VARIANT XSCID LYS-68.  
 RX MEDLINE; 94375038.  
 RA Markiewicz S., Subtil A., Dautry-Varsat A., Fischer A.,  
 RA de Saint Basile G.;  
 RT "Detection of three nonsense mutations and one missense mutation in  
 RT the interleukin-2 receptor gamma chain gene in SCIDX1 that  
 RT differently affect the mRNA processing.";  
 RL Genomics 21:291-293(1994).  
 RN [10]  
 RN VARIANT XSCID HIS-162.  
 RX MEDLINE; 94300093.  
 RA Ishii N., Asao H., Kimura Y., Takeshita T., Nakamura M., Tsuchiya S.,  
 RA Kono T., Maeda M., Uchiyama T., Sugamura K.;  
 RT "Impairment of ligand binding and growth signaling of mutant IL-2  
 RT receptor gamma-chains in patients with X-linked severe combined  
 RT immunodeficiency.";  
 RL J. Immunol. 153:1310-1317(1994).  
 RN [11]  
 RN VARIANT XSCID ASN-39.  
 RX MEDLINE; 95023932.  
 RA Disanto J.P., Rioux-Laucat F., Dautry-Varsat A., Fischer A.,  
 RA de Saint Basile G.;  
 RT "Defective human interleukin 2 receptor gamma chain in an atypical X

RT chromosome-linked severe combined immunodeficiency with peripheral T  
 RT cells.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:9466-9470(1994).  
 RN [12]  
 RN VARIANTS XSCID CYS-226 AND HIS-226.  
 RX MEDLINE; 95397841.  
 RA Pepper A.E., Buckley R.H., Small T.N., Puck J.M.;  
 RT "Two mutational hotspots in the interleukin-2 receptor gamma chain  
 RT gene causing human X-linked severe combined immunodeficiency.";  
 RL Am. J. Hum. Genet. 57:564-571(1995).  
 RN [13]  
 RN VARIANT XSCID SER-183.  
 RX MEDLINE; 96013903.  
 RA Clark P.A., Lester T., Genet S., Jones A.M., Hendriks R.,  
 RA Levisky R.L., Kinnon C.;  
 RT "Screening for mutations causing X-linked severe combined  
 RT immunodeficiency in the IL-2R gamma chain gene by single-strand  
 RT conformation polymorphism analysis.";  
 RL Hum. Genet. 96:427-432(1995).  
 RN [14]  
 RN VARIANT XSCID GLN-HIS-TRP INS-237.  
 RX MEDLINE; 95164726.  
 RA Puck J.M., Pepper A.E., Bedard P.-M., Laframboise R.;  
 RT "Female germ line mosaicism as the origin of a unique IL-2 receptor  
 RT gamma-chain mutation causing X-linked severe combined  
 RT immunodeficiency.";  
 RL J. Clin. Invest. 95:895-899(1995).  
 RN [15]  
 RN VARIANT XSCID GLN-271.  
 RX MEDLINE; 95190013.  
 RA Schmalstieg F.C., Leonard W.J., Noguchi M., Berg M., Rudloff H.E.,  
 RA Deane R.M., Dave S.K., Brooks E.G., Goldman A.S.;  
 RT "Massive mutation in exon 7 of the common gamma chain gene causes a  
 RT moderate form of X-linked combined immunodeficiency.";  
 RL J. Clin. Invest. 95:1169-1173(1995).  
 RN [16]  
 RN VARIANT XSCID ARG-115.  
 RX MEDLINE; 97042245.  
 RA Stephan V., Wahn V., le Deist F., Dirksen U., Broker B.,  
 RA Muller-Fleckenstein I., Horneff G., Schroten H., Fischer A.,  
 RA de Saint Basile G.;  
 RT "Atypical X-linked severe combined immunodeficiency due to possible  
 RT spontaneous reversion of the genetic defect in T cells.";  
 RL New Engl. J. Med. 335:1563-1567(1996).  
 RN [17]  
 RN VARIANT XSCID GLN-285.  
 RX MEDLINE; 97295088.  
 RA Jones A.M., Clark P.A., Katz F., Genet S., McMahon C., Alterman L.,  
 RA Cant A., Kinnon C.;  
 RT "B-cell-negative severe combined immunodeficiency associated with a  
 RT common gamma chain mutation.";  
 RL Hum. Genet. 99:677-680(1997).  
 RN [18]  
 RN VARIANT XSCID CYS-222.  
 RX MEDLINE; 98084061.  
 RA Sharfe N., Shahar M., Roifman C.M.;  
 RT "An interleukin-2 receptor gamma chain mutation with normal thymus  
 RT morphology.";  
 RL J. Clin. Invest. 100:3036-3043(1997).  
 CC -!- FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF  
 CC INTERLEUKINS.  
 CC -!- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND  
 CC PROBABLY ALSO THE IL-13 RECEPTORS.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- DISEASE: DEFECTS IN IL2RG ARE THE CAUSE OF A SEVERE COMBINED  
 CC IMMUNODEFICIENCY, WHICH IS KNOWN AS AGAMAGLOBULINEMIA, SWISS TYPE  
 CC OR X-LINKED SEVERE COMBINED IMMUNODEFICIENCY DISEASE (XSCID).  
 CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD132 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd132.htm".  
 CC -!- DATABASE: NAME=IL2RGbase; NOTE=X-linked SCID mutation database;  
 CC WWW="http://www.nhgri.nih.gov/DIR/IL2RGbase.html".

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D11086; BAA01857.1; -
DR EMBL; L12183; AAA59145.1; -
DR EMBL; L12178; AAA59145.1; JOINED.
DR EMBL; L12176; AAA59145.1; JOINED.
DR EMBL; L12177; AAA59145.1; JOINED.
DR EMBL; L12179; AAA59145.1; JOINED.
DR EMBL; L12180; AAA59145.1; JOINED.
DR EMBL; L12181; AAA59145.1; JOINED.
DR EMBL; L12182; AAA59145.1; JOINED.
DR EMBL; L19546; AAC37524.1; -
DR PIR; A42565; A42565.
DR PDB; 1ILM; 26-JAN-95.
DR PDB; 1ILN; 26-JAN-95.
DR MIM; 300400; -
DR MIM; 308380; -
DR PROSITE; PS00241; RECEPTOR_CYTOKINES_1; 1.
DR PROSITE; PS00340; RECEPTOR_CYTOKINES_2; FALSE_NEG.
DR PFAM; PF00041; fn3; 1.
KW Receptor; Transmembrane; Glycoprotein; Signal; Disease mutation;
...
Note: remainder of annotations omitted.
Query Match 7.0%; Score 195; DB 1; Length 369;
Best Local Similarity 25.9%; Pred. No. 1.28e-16;
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;
Db 59 EVQCFVNEVMNCTNNSSEPPQTNLTHVYKNSNDKVKQKSHYFSEITSGCOIQ 118
QY 31 QIQIYFNLETQVTVNASKYSR-TNLTFFHYR-NGD-EAYDQCTNYLQEGHTSGCLLD 87
Db 119 KKEHLQYTFVVLQDPRPRQATQMLKQLNVLIPWAPENLTLHLKSESOLELNNRNF 178
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYIL-KPSPFKHVR-FWQDQAVTVCSDL 142
Db 179 LNHCLHLVQYRVDWDSWTSQSYDYRHKFSLPSVDGQKRYTFVRSR-FNPLOGSAOH 237
QY 143 YGD-LLYEVOYRSPFDTEW-QSKOENTCNVTEGLDAEKYSFVVRVAMEDVYGPDTYP 200
Db 238 SEWSHPHWH 246
QY 201 SDWSEVTCW 209
RESULT 4
ID CYRG_MOUSE STANDARD; PRT; 369 AA.
AC P34902;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C)
DE (INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64).
GN IL2RG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93277575.
RA Kumaki S., Kondo M., Takeshita T., Asao H., Nakamura M., Sugamura K.;
RT "Cloning of the mouse interleukin 2 receptor gamma chain:
RT demonstration of functional differences between the mouse and human
RT receptors."
RL Biochem. Biophys. Res. Commun. 193:356-363(1993).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=CBA/CA;
RX MEDLINE; 93391374.
RA Cao X., Kozak C.A., Liu Y.J., Noguchi M., O'Connell E., Leonard W.J.;
RT "Characterization of cDNAs encoding the murine interleukin 2 receptor
RT (IL-2R) gamma chain: chromosomal mapping and tissue specificity of
RT IL-2R gamma chain expression."
RL Proc. Natl. Acad. Sci. U.S.A. 90:8464-8468(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93366191.
RA Kobayashi N., Nakagawa S., Minami Y., Taniguchi T., Kono T.;
RT "Cloning and sequencing of the cDNA encoding a mouse IL-2 receptor
RT gamma."
RL Gene 130:303-304(1993).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95104285.
RA Disanto J.P., Certain S., Wilson A., Macdonald H.R., Avner P.,
RA Fischer A., de Saint Basile G.;
RT "The murine interleukin-2 receptor gamma chain gene: organization,
RT chromosomal localization and expression in the adult thymus."
RL Eur. J. Immunol. 24:3014-3018(1994).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=B6.S;
RX MEDLINE; 96341745.
RA Chiu R.K., Droll A., Cooper D.L., Dougherty S.T., Dirks J.F.,
RA Dougherty G.J.;
RT "Molecular mechanisms regulating the hyaluronan binding activity of
RT the adhesion protein CD44."
RL J. Neurocol. 26:231-239(1995).
CC -!- FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF
CC INTERLEUKINS.
CC -!- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND
CC PROBABLY ALSO THE IL-13 RECEPTORS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D13821; BAA02974.1; -
DR EMBL; U21795; AAA64279.1; -
DR EMBL; D13565; BAA02760.1; -
DR EMBL; L20048; AAA39286.1; -
DR EMBL; S75852; AAB32904.1; -
DR EMBL; S75844; AAB32904.1; JOINED.
DR EMBL; S75845; AAB32904.1; JOINED.
DR EMBL; S75847; AAB32904.1; JOINED.
DR EMBL; S75848; AAB32904.1; JOINED.
DR EMBL; S75849; AAB32904.1; JOINED.
DR EMBL; S75850; AAB32904.1; JOINED.
DR EMBL; S75851; AAB32904.1; JOINED.
DR EMBL; X75337; CA853085.1; -
DR PIR; JN0592; JN0592.
DR PIR; JN0775; JN0775.
DR HSSP; P31785; 1ILN.
DR MGD; MGI:96551; IL2RG.
DR PROSITE; PS00241; RECEPTOR_CYTOKINES_1; 1.
DR PROSITE; PS00340; RECEPTOR_CYTOKINES_2; FALSE_NEG.
DR PFAM; PF00041; fn3; 1.
KW Receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1..22 BY SIMILARITY.
FT CHAIN 23..369 CYTOKINE RECEPTOR COMMON GAMMA CHAIN.
FT DOMAIN 23..263 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 264..284 POTENTIAL.

```





[illegible]

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

CC EMBL; D13566; BAA02761.1; -  
CC PIR; A46713; A46713.  
CC HSP; P19235; IERP.  
CC PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
CC PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; 1.  
CC PFAM; PF00041; fn3; 1.  
CC KW Receptor; Transmembrane; Glycoprotein; Signal.  
CC SIGNAL ; 1 24  
CC CHAIN ; 25 507 BY SIMILARITY  
CC DOMAIN ; 25 249 ERYTHROPOIETIN RECEPTOR.  
CC TRANSMEM ; 250 272 EXTRACELLULAR (POTENTIAL).  
CC DOMAIN ; 273 507 POTENTIAL.  
CC DOMAIN \*147 212 CYTOPLASMIC (POTENTIAL).  
CC DISULFID ; 52 62 FIBONECTIN TYPE-III.  
CC DISULFID 90 106 BY SIMILARITY.  
CC CARBOHYD 75 75 BY SIMILARITY.  
CC SEQUENCE 507 AA; 55499 MW; AC79AF22D06A7312 CRC64;

Query Match 5.9%; Score 164; DB 1; Length 507;  
Best Local Similarity 45.2%; Pred. No. 5.49e-11;  
Matches 28; Conservative 11; Mismatches 19; Indels 4; Gaps 3;

247 LDPILITLSL-ILVLISLILITVALLSHRRALRQKINPGISPENFGLFTTHKGNFQL 305  
| : | | | | | | | | | | : : : : | : | | | | | | | | | |  
229 LSKFLISSLAILMLVSLLL--LSLWLRLVRKKFLI-PSVPDPKSIFFGLFEIHQGNFQE 285  
| : | | | | | | | | | | : : : : | : | | | | | | | | | |  
306 WL 307  
| :  
286 WI 287

RESULT 8  
ID EPOR\_MOUSE STANDARD; PRT; 507 AA.  
AC P14753; Q63852;  
DT 01-APR-1990 (Rel. 14, Created)  
DD 01-APR-1990 (Rel. 14, Last sequence update)  
DE 01-NOV-1995 (Rel. 32, Last annotation update)  
GN ERYTHROPOIETIN RECEPTOR PRECURSOR (EPO-R).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OT Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
[1]  
RP SEQUENCE FROM N.A.  
SA MEDLINE; 89195238.  
SA D'Andrea A.B., Lodish H.F., Wong G.G.;  
SA "Expression cloning of the murine erythropoietin receptor.";  
SA Cell 57:277-285(1989).  
[2]  
RP SEQUENCE FROM N.A.  
SA STRAIN-BALB/C; TISSUE=LIVER;  
SA MEDLINE; 91080149.  
SA Kuramochi S., Ikawa Y., Todokoro K.;  
SA "Characterization of murine erythropoietin receptor genes.";  
SA J. Mol. Biol. 216:567-575(1990).  
[3]  
RP SEQUENCE FROM N.A.  
SA MEDLINE; 92017832.  
SA Hino M., Tojo A., Misawa Y., Morii H., Takaku F., Shibuya M.;  
SA "Unregulated expression of the erythropoietin receptor gene caused by  
SA insertion of spleen focus-forming virus long terminal repeat in a  
SA murine erythroleukemia cell line.";  
SA Mol. Cell. Biol. 11:5527-5533(1991).  
[4]  
RP SEQUENCE OF 1-27 FROM N.A.  
SA YOUSSEFIAH H., Zou L., Orlin S.H., D'Andrea A.B., Lodish H.F.,  
SA MEDLINE; 90287158.

CC  
CC  
CC

Thu May 11 06:49:25 2000

```

RT "Structure and transcription of the mouse erythropoietin receptor
RL gene.";
RN Mol. Cell. Biol. 10:3675-3682(1990).
RN [5]
RN SEQUENCE OF 1-24 FROM N.A.
RX MEDLINE; 91201346.
RA Lacombe C., Chretien S., Lemarchandel V., Mayeux P., Romeo P.H.,
RA Gisselbrecht S., Cartton J.P.;
RA "Spleen focus-forming virus long terminal repeat insertional
RT activation of the murine erythropoietin receptor gene in the t3c1-2
RT friend leukemia cell line.";
RL J. Biol. Chem. 266:6952-6956(1991).
RN [6]
RN MUTAGENESIS.
RX MEDLINE; 93180826.
RA Miura O., Cleveland J.L., Ihle J.N.;
RT "Inactivation of erythropoietin receptor function by point mutations
RT in a region having homology with other cytokine receptors.";
RL Mol. Cell. Biol. 13:1788-1795(1993).
CC -1- FUNCTION: RECEPTOR FOR ERYTHROPOIETIN. MAY PLAY A ROLE IN THE
CC MECHANISM OF ERYTHROPOIETIN-INDUCED ERYTHROBLAST PROLIFERATION
CC AND DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J04843; AAA37571.1; .
DR EMBL; X53081; CAA37248.1; .
DR EMBL; M38133; AAA37572.1; .
DR EMBL; M62360; AAA37582.1; .
DR EMBL; S59388; CAB31799.1; .
DR EMBL; A32385; A32385.
DR PIR; A41686; A41686.
DR PIR; S13249; S13249.
DR PIR; S14081; S14081.
DR HSP; P19235; IEBP.
DR MGD; MGI:95408; EPOR.
DR PROSITE; PS00241; RECEPTOR_CYTOKINES_1; 1.
DR PROSITE; PS00340; RECEPTOR_CYTOKINES_2; 1.
DR PFAM; PF00041; fn3; 1.
KW Receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 507
FT DOMAIN 25 249
FT TRANSMEM 250 272
FT DOMAIN 273 507
FT DOMAIN 147 212
FT DISULFID 52 62
FT DISULFID 90 106
FT CARBOHYD 75 75
FT CONFLICT 291 291
FT SEQUENCE 507 AA; 55194 MW; 067657A2E26451CA CRC64;
Query Match 5.88; Score 163; DB 1; Length 507;
Best Local Similarity 46.8%; Pred. No. 8.22e-11;
Matches 29; Conservative 9; Mismatches 20; Indels 4; Gaps 3;
Db 247 LDPLILTLSS-ILVLISLLTLVIALSHRRTLOQKINWGPISPESEFGLTHKGNFQL 305
QY 229 LSKFILISSLAIIIMVSLLL--LSLWLKRWVKFLI-PSVPDPKSIFFGLFEIHQGNFQE 285
Db 306 WL 307
QY 286 WI 287

```

```

RESULT 9
ID EPOR_HUMAN STANDARD; PRT; 508 AA.
AC P19235;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ERYTHROPOIETIN RECEPTOR PRECURSOR (EPO-R).
GN EPOR
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 91372359.
RA Ehrenman K., St John T.;
RT "The erythropoietin receptor gene: cloning and identification of
RT multiple transcripts in an erythroid cell line OCIML.";
RL Exp. Hematol. 19:973-977(1991).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE; 90304340.
RA Jones S.S., D'Andrea A., Haines L.L., Wong G.G.;
RT "Human erythropoietin receptor: cloning, expression, and biologic
RT characterization.";
RL Blood 76:31-35(1990).
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE; 92399733.
RA Noguichi C.T., Bae K.S., Chin K., Wada Y., Schechter A.N.,
RA Hawkins W.D.;
RT "Cloning of the human erythropoietin receptor gene.";
RL Blood 78:2548-2556(1991).
RN [4]
RN SEQUENCE OF 1-96 FROM N.A.
RX TISSUE-PLACENTA;
RX MEDLINE; 92399734.
RA Macouche L., Tournamille C., Hattab C., Boffa G., Cartton J.P.,
RA Chretien S.;
RT "Cloning of the gene encoding the human erythropoietin receptor.";
RL Blood 78:2557-2563(1991).
RN [5]
RN SEQUENCE OF 1-17 FROM N.A.
RX MEDLINE; 92147143.
RA Penny L.A., Forget B.G.;
RT "Genomic organization of the human erythropoietin receptor gene.";
RL Genomics 11:974-980(1991).
RN [6]
RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 34-244.
RX MEDLINE; 96291992.
RA Livnah O., Stura E.A., Johnson D.L., Middleton S.A., Mulcahy L.S.,
RA Livnah O., Stura E.A., Johnson D.L., Middleton S.A., Mulcahy L.S.,
RA Wright N.C., Dower W.J., Jolliffe L.K., Wilson I.A.;
RT "Functional mimicry of a protein hormone by a peptide agonist: the
RT EPO receptor complex at 2.8 A.";
RL Science 273:464-471(1996).
RN [7]
RN X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 34-244.
RX MEDLINE; 99023198.
RA Livnah O., Johnson D.L., Stura E.A., Farrell F.X., Barbone F.P.,
RA You Y., Liu K.D., Goldsmith M.A., He W., Krause C.D., Pestka S.,
RA Jolliffe L.K., Wilson I.A.;
RT "An antagonist peptide-EPO receptor complex suggests that receptor
RT dimerization is not sufficient for activation.";
RL Nat. Struct. Biol. 5:993-1004(1998).
CC -1- FUNCTION: RECEPTOR FOR ERYTHROPOIETIN. MAY PLAY A ROLE IN THE
CC MECHANISM OF ERYTHROPOIETIN-INDUCED ERYTHROBLAST PROLIFERATION
CC AND DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; M34986; AA52401.1; -  
 CC EMBL; M60459; AA52403.1; -  
 CC EMBL; S45332; AAB23271.1; -  
 CC EMBL; M76595; AA52393.1; -  
 CC EMBL; M77244; AA52392.1; -  
 CC PIR; A43799; A43799; -  
 CC PIR; A49824; A49824; -  
 CC PIR; A53958; A53958; -  
 CC PDB; 1EBP; 29-JUL-97.  
 CC PDB; 1EBA; 18-NOV-98.  
 CC MIM; 133171; -

DR PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
 DR PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; 1.  
 DR PFAM; PF00041; fn3; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Signal; 3D-structure.

FT SIGNAL 1 24  
 FT CHAIN 25 508  
 FT DOMAIN 25 508  
 FT TRANSMEM 251 273  
 FT DOMAIN 274 508  
 FT DOMAIN 148 213  
 FT DISULFID 52 62  
 FT DISULFID 91 107  
 FT CARBOHYD 76 76  
 SQ SEQUENCE 508 AA; 55065 MW; F9F326E162E9512A CRC64;

Query Match 5.7%; Score 160; DB 1; Length 508;  
 Best Local Similarity 42.6%; Pred. No. 2.75e-10;  
 Matches 26; Conservative 12; Mismatches 21; Indels 2; Gaps 2;

Db 248 LDPILTLIIIVLIVLTLVLLSHLRALKOKTWPGIPSPSEFEGFTTHKGNFQW 307  
 QY 229 LSKFLLISLALLMVSL-LLSLKWLKVRKFLI-PSVPDPKRSIFPGFIEHOGNFQW 286

Db 308 L 308

QY 287 I 287

RESULT 10  
 ID CP23 MOUSE STANDARD; PRT; 490 AA.

AC P56655;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CYTOCHROME P450 2C38 (EC 1.14.14.1) (CYF11C38).  
 GN CYP2C38.  
 OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 CC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-CD-1.

RX MEDLINE; 98389577.

RA Luo G., Zeldin D.C., Blaisdell J.A., Hodgson E., Goldstein J.A.;

RT "Cloning and expression of murine CYP2C8 and their ability to

metabolize arachidonic acid.";

RL Arch. Biochem. Biophys. 357:45-57(1998).

CC -1- FUNCTION: METABOLIZES ARACHIDONIC ACID TO PRODUCE 11,12-EET.

CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +

OXIDIZED FLAVOPROTEIN + H(2)O.

CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.

CC -1- TISSUE SPECIFICITY: LIVER, BRAIN, KIDNEY, AND INTESTINE, WITH

TRACE AMOUNTS IN LUNG AND HEART.

CC -1- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER

CC AND CARCINOGENS.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF047725; AAD13720.1; -  
 CC MGD; MGI:1306819; CYP2C38.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT BINDING 435 435  
 FT HEME (BY SIMILARITY).  
 SQ SEQUENCE 490 AA; 56089 MW; C2922E7EC36A410C CRC64;

Query Match 4.5%; Score 126; DB 1; Length 490;  
 Best Local Similarity 33.9%; Pred. No. 1.16e-04;  
 Matches 19; Conservative 18; Mismatches 17; Indels 2; Gaps 2;

Db 3 LVTFL-~~MT~~LSLLSLWRQ-RSRGRPLPGPTPTPIIGNFLQIDVKNFQSLTN 56  
 QY 234 LLSLALLMVSL-LLSLKWLKVRKFLIPSPDKSIFPGFIEHOGNFQWITD 289

RESULT 11  
 ID IL13L MOUSE STANDARD; PRT; 424 AA.

AC O09030;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE INTERLEUKIN-13 RECEPTOR ALPHA-1 CHAIN PRECURSOR (IL-13R-ALPHA-1) (IL-13RA-1) (INTERLEUKIN-13 BINDING PROTEIN) (NR4).  
 GN IL13RAL OR IL13RA OR IL13R.  
 OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 CC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96133964.  
 RA Hilton D.J., Zhang J.-G., Metcalf D., Alexander W.S., Nicola N.A.,  
 RA Willson T.A.;

RT "Cloning and characterization of a binding subunit of the interleukin  
 RT 13 receptor that is also a component of the interleukin 4 receptor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:497-501(1996).

CC -1- FUNCTION: BINDS IL-13 WITH A LOW AFFINITY. TOGETHER WITH IL-4R-  
 CC ALPHA CAN FORM A FUNCTIONAL RECEPTOR FOR IL-13. ALSO SERVES AS AN  
 CC ALTERNATE ACCESSORY PROTEIN TO THE COMMON CYTOKINE RECEPTOR GAMMA  
 CC CHAIN FOR IL-4 SIGNALING, BUT CANNOT REPLACE THE FUNCTION OF  
 CC GAMMA-C IN ALLOWING ENHANCED IL-2 BINDING ACTIVITY (BY  
 CC SIMILARITY).

CC -1- SUBUNIT: INTERLEUKIN-13 RECEPTOR IS A COMPLEX OF IL4R-ALPHA,  
 CC IL13R-ALPHA, AND POSSIBLY OTHER COMPONENTS (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -1- TISSUE SPECIFICITY: SPLEEN, LIVER, THYMUS, HEART, LUNG, KIDNEY,  
 CC TESTIS, STOMACH, BRAIN, SKIN, AND COLON; BUT NOT SKELETAL MUSCLE.

CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; S80963; AAB50695.1; -  
 CC MGD; MGI:105052; IL13RA.  
 DR Receptor; Transmembrane; Glycoprotein; Immunoglobulin domain; Signal.

Thu May 11 06:49:25 2000

FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 424 INTERLEUKIN-13 RECEPTOR ALPHA-1 CHAIN.  
FT DOMAIN 26 340 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 341 360 POTENTIAL.  
FT DOMAIN 365 424 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 37 100 IG-LIKE C2-TYPE DOMAIN.  
FT DISULFID 44 93 POTENTIAL.  
FT DISULFID 132 142 BY SIMILARITY.  
FT DISULFID 171 183 BY SIMILARITY.  
FT CARBOHYD 35 35 POTENTIAL.  
FT CARBOHYD 59 59 POTENTIAL.  
FT CARBOHYD 103 103 POTENTIAL.  
FT CARBOHYD 136 136 POTENTIAL.  
FT CARBOHYD 262 262 POTENTIAL.  
FT CARBOHYD 338 338 POTENTIAL.  
SQ SEQUENCE 424 AA; EB8330A0DC82C9F9 CRC64;  
Query Match 4.3%; Score 119; DB 1; Length 424;  
Best Local Similarity 20.4%; Pred. No. 1.37e-03;  
Matches 58; Conservative 88; Mismatches 110; Indels 29; Gaps 26;  
Db 120 EGDPEAVTELKCIWNLSYMKSCWLPGRWTSPTDHYLYWYSSLSKRCQEN-IVREG 178  
Qy 23 QGGAAEGV-QIIFYNLETQVVTNASKY-SR-TNLTFFHYRFGNGDEAYDOCTNYLLQEG 79  
Db 179 QHICASFKLKVERSEFQHNQVQIMVKNAGKIRPSCKIVSLTSYKDPDP-HIKHLLKN 237  
Qy 80 HTSGC-L-LD-AEQR-D-DILYFSIRNGTHPVFTASR--WVYLYLKSSPKHVR-FSWHQ 131  
Db 238 GALLYQWKNPNFRSCLTYEVEVNNNT-OTDRHNLVEEDKCONSDRMWGSTSCFQL 296  
Qy 132 DAVTVTCSD-LSY-GD-LLYEVOVSPEDTEWQS--K-OENTC-NVTIE-GLDAEKCYSF 183  
Db 297 -PGVLA-DAVTVRY-RVK-TNKLCDNKLMSDSEWSEKQSIGKQNSFTYTMLLTPIVF 352  
Qy 184 WVRKAMEDVYGPDPYPSDWSEVTCQWGEINDACAETPTPKPKLSKF--ILISSAIL 241  
Db 353 VAVAVIIL-LFYLKLLKIIPPIPGKIFKEMFGDQNDTLHW 396  
Qy 242 LNVSLLSLWKLVRKFLIPSPDKSIFPGLFEIHQGNFOEW 286  
RESULT 12 STANDARD; PRT; 490 AA.  
ID CPZ4\_MOUSE  
AC P56656;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME P450 2C39 (EC 1.14.14.1) (CYFIC39).  
GN CPZC39.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN  
RP STRAIN=CD-1;  
RC SEQUENCE FROM N.A.  
RX MEDLINE; 98389577.  
RA Luo G., Zeldin D.C., Blaisdell J.A., Hodgson E., Goldstein J.A.;  
RT "Cloning and expression of murine CYP2C3 and their ability to  
metabolize arachidonic acid."  
RL Arch. Biochem. Biophys. 357:45-57(1998).  
CC -1- FUNCTION: METABOLIZES ARACHIDONIC ACID TO PRODUCE 14,15-CIS-  
EPOXYEICOSATRIENOIC ACID (BET).  
CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) -> ROH +  
OXIDIZED FLAVOPROTEIN + H(2)O.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
CC -1- TISSUE SPECIFICITY: LIVER.  
CC -1- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER  
TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES,  
AND CARCINOGENS.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC This-SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; AF047726; AAD13721.1;  
DR MGD; MGI:1305818; CYP2C39.  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum.  
FT BINDING 435 435 HEME (BY SIMILARITY).  
SQ SEQUENCE 490 AA; 55902 MW; 97EF56D0A4B728C3 CRC64;  
Query Match 4.1%; Score 113; DB 1; Length 490;  
Best Local Similarity 33.9%; Pred. No. 1.05e-02;  
Matches 19; Conservative 16; Mismatches 19; Indels 2; Gaps 2;  
Db 3 LVTFLLVTLSLLSLWROSCGRGSL-PPGPTFPPIGFLQIDMKNFQSGLTN 56  
Qy 234 LISSAILMVSLLSLWKLVRKFLIPSPDKSIFPGLFEIHQGNFOEWIID 289  
RESULT 13 STANDARD; PRT; 878 AA.  
ID IL3B\_MOUSE  
AC P26954;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 01-FEB-2000 (Rel. 39, Last annotation update)  
DE INTERLEUKIN-3 RECEPTOR CLASS II BETA CHAIN PRECURSOR (COLONY  
STIMULATING FACTOR 2 RECEPTOR, BETA 2 CHAIN).  
GN CSF2RB2 OR AL2CA OR IL3RB2 OR IL3R.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90117145.  
RA Itoh N., Yonehara S., Schreurs J., Gorman D.M., Maruyama K., Ishii A.,  
RA Yahara I., Arai K., Miyajima A.;  
RT "Cloning of an interleukin-3 receptor gene: a member of a distinct  
receptor gene family."  
RL Science 247:324-327(1990).  
CC -1- FUNCTION: IN MOUSE THERE ARE TWO CLASSES OF HIGH-AFFINITY IL-3  
RECEPTORS. ONE CONTAINS THIS IL-3-SPECIFIC BETA CHAIN AND THE  
OTHER CONTAINS THE BETA CHAIN ALSO SHARED BY HIGH-AFFINITY IL-5  
AND GM-CSF RECEPTORS.  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; M29855; AAA39295.1;  
DR PIR; A40091; A40091.  
DR MGD; MGI:1339760; CSF2RB2.  
DR PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
DR PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; 1.  
DR PFAM; PF00041; fn3; 2.  
KW Receptor; Transmembrane; Glycoprotein; Signal.  
FT SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 878 INTERLEUKIN-3 RECEPTOR CLASS II BETA  
DOMAIN 23 440 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 441 462 POTENTIAL.  
FT DOMAIN 463 878 CYTOPLASMIC (POTENTIAL).

```

FT DISULFID 39 49 BY SIMILARITY.
FT DISULFID 78 95 BY SIMILARITY.
FT DISULFID 254 264 BY SIMILARITY.
FT DISULFID 293 310 BY SIMILARITY.
FT CARBOHYD 62 62 POTENTIAL.
FT CARBOHYD 350 350 POTENTIAL.
SQ SEQUENCE 878 AA; 97195 MW; 8BEC9092ADC24D56 CRC64;

Query Match 4.1%; Score 115; DB 1; Length 878;
Best Local Similarity 32.4%; Pred. No. 5.37e-03;
Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 370 YIDHTFQVQYKKKSES-WKDSKTENLGRVNSMDLPQLPDTSCYCARVKRPSD-YD-GI 426
QY 143 YGDLLEVOYRSPDTEW-QSKQENTCNV-TIE-G-LDAEKCYCFWVRKAMEDVYGPDT 198
Db 427 W-SEWSNYTW 436
QY 199 YPSDMSVTCW 209

RESULT 14
ID CYRB-HUMAN STANDARD; PRT; 897 AA.
AC P32927;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CYTOKINE RECEPTOR COMMON BETA CHAIN PRECURSOR (CD131 ANTIGEN).
GN CSF2RB OR IL5RB OR IL3RB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91088571.
RA Hayashida K., Kitamura T., Gorman D.M., Arai K., Yokota T.,
RA Miyajima A.;
RT "Molecular cloning of a second subunit of the receptor for human
RT granulocyte-macrophage colony-stimulating factor (GM-CSF):
RT reconstitution of a high-affinity GM-CSF receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:9655-9659(1990).
RN [2]
RP REVISION TO 454.
RA Kitamura T.;
RA Submitted (FEB-1991) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: HIGH AFFINITY RECEPTOR FOR INTERLEUKIN-3, INTERLEUKIN-5
CC AND GRANULOCYTE-MACROPHAGE COLONY-STIMULATING FACTOR.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. THE BETA
CC CHAIN IS COMMON TO THE IL-3, IL-5 AND GM-CSF RECEPTORS.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide Cw131 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cdw131.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M59941; AAA18171.1;
CC PIR; A39255; A39255.
CC
CC MIM; 138981;
CC
CC PROSITE; PS00241; RECEPTOR_CYTOKINES_1; 2.
CC PROSITE; PS00340; RECEPTOR_CYTOKINES_2; 1.
CC PFAM; PF00041; fn3; 2.
CC
CC Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 16 POTENTIAL
FT CHAIN 17 897 CYTOKINE RECEPTOR COMMON BETA CHAIN.
FT DOMAIN 17 443 EXTRACELLULAR (POTENTIAL).

```

```

FT TRANSMEM 444 460 POTENTIAL.
FT DOMAIN 461 897 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 129 238 FIBRONECTIN TYPE-III.
FT DOMAIN 336 434 FIBRONECTIN TYPE-III.
FT DISULFID 35 45 BY SIMILARITY.
FT DISULFID 75 91 BY SIMILARITY.
FT CARBOHYD 58 58 POTENTIAL.
FT CARBOHYD 191 191 POTENTIAL.
FT CARBOHYD 346 346 POTENTIAL.
SQ SEQUENCE 897 AA; 97335 MW; 3398E37FDB8F393A CRC64;

```

Query Match 4.1%; Score 113; DB 1; Length 897;  
 Best Local Similarity 31.9%; Pred. No. 1.05e-02;  
 Matches 22; Conservative 10; Mismatches 34; Indels 3; Gaps 3;

```

Db 372 FEIQYRKDTAT-WKDSKTETLQNAHSMALPAPSTPSTVWARVVRTSRTGVNGIWSWSE 430
QY 148 YEVOYRSPDTEW-QSKQENTCNVTIEGLDA-EKCYFVVRKAMEDVYGPDTYPSDWSE 205
Db 431 ARSWDTESV 439
QY 206 VTCWQGEI 214

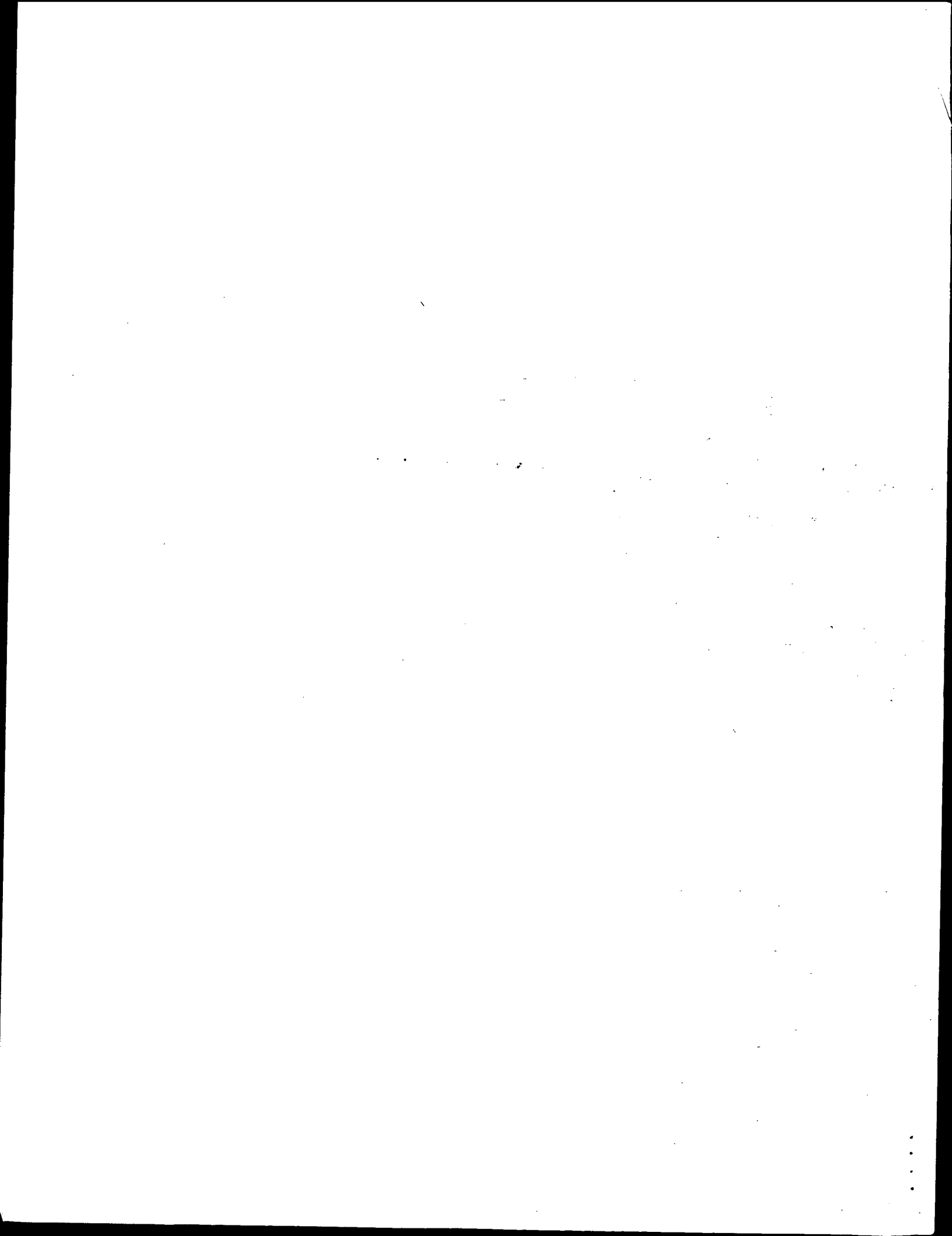
```

```

RESULT 15
ID CPC7-RAT STANDARD; PRT; 490 AA.
AC P05179;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CYTOCHROME P450 2C7 (EC 1.14.14.1) (CYP11C7) (P450F) (PTF1).
GN CYP2C7 OR CYP2C-7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=LIVER;
RX MEDLINE; 88087187.
RA Kimura H., Yoshioka H., Sogawa K., Sakai Y., Fujii-Kuriyama Y.;
RT "Complementary DNA cloning of cytochrome P-450s related to P-450(M-1)
RT from the complementary DNA library of female rat livers. Predicted
RT primary structures for P-450F, PB-1, and PB-1-related protein with a
RT bizarre replacement block and their mode of transcriptional
RT expression.";
RL J. Biol. Chem. 263:701-707(1988).
RN [2]
RP SEQUENCE OF 8-490 FROM N.A.
RX MEDLINE; 86278140.
RA Gonzalez F.J., Kimura S., Song B.-J., Pastewka J., Gelboin H.V.,
RA Hardwick J.P.;
RT "Sequence of two related P-450 mRNAs transcriptionally increased
RT during rat development. An R.dre.1 sequence occupies the complete 3'
RT untranslated region of a liver mRNA.";
RL J. Biol. Chem. 261:10667-10672(1986).
RN [3]
RP SEQUENCE OF 87-490 FROM N.A.
RX MEDLINE; 87101095.
RA Friedberg T., Waxman D.J., Atchison M., Kumar A., Haaparanta T.,
RA Raphael C., Adesnik M.;
RT "Isolation and characterization of cDNA clones for cytochromes P-450
RT immunohistochemically related to rat hepatic P-450 form PB-1.";
RL Biochemistry 25:7975-7983(1986).
CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC ACIDS, AND XENOBIOTICS.
CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) -> ROH +
CC OXIDIZED FLAVOPROTEIN + H(2)O.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
CC -!- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER
CC TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES,

```







\*\*\*\*\*  
M P S R L H  
(TW)  
\*\*\*\*\*

Release 3.1A John F. Collins, Blocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution Rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:10:01 2000; MasPar time 13.95 Seconds  
629.947 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-376-430-2  
Description: (1-371) from US09376430A.pap (1 of 25)  
Perfect Score: 2788  
Sequence: 1 MGRLLVLLGAAVFLGGMW.....DVTIGGTFVMDRSYVAL 371

Scoring table: PAM 150  
Gap 11

Searched: 189963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq35  
1:geneseqp

Statistics: Mean 35.238; Variance 152.691; scale 0.231

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	200	7.2	482	1 W31646	Human cytokine recepto	2.25e-07
2	195	7.0	230	1 R22934	Interleukin 4 componen	5.99e-07
3	195	7.0	230	1 R47151	IL-2 receptor gamma ch	5.99e-07
4	195	7.0	230	1 R47150	IL-2 receptor gamma ch	5.99e-07
5	195	7.0	347	1 R47149	IL-2 receptor gamma ch	5.99e-07
6	195	7.0	369	1 R47148	IL-2 receptor gamma ch	5.99e-07
7	188	6.7	369	1 R59094	Murine IL-2R gamma.	7.34e-06
8	176	6.5	500	1 W64057	Human IL-9 receptor pr	2.36e-05
9	176	6.3	501	1 W64055	Human IL-9 receptor pr	2.36e-05
10	175	6.3	501	1 W64056	Human IL-9 receptor pr	2.36e-05
11	163	5.8	507	1 R47517	MEL EPO receptor.	2.79e-04
12	163	5.8	507	1 R69502	Mouse erythropoietin r	2.79e-04
13	163	5.8	507	1 R50327	Mouse soluble EPO rece	2.79e-04
14	160	5.7	508	1 R06511	EPO receptor sequence	4.90e-04
15	160	5.7	508	1 R69503	Human erythropoietin r	4.90e-04
16	160	5.7	508	1 R06512	EPO receptor.	4.90e-04
17	160	5.7	508	1 R70032	Human erythropoietin r	4.90e-04
18	160	5.7	508	1 R47518	Human EPO receptor.	4.90e-04
19	132	4.7	383	1 W35294	Murine IL-13 binding c	8.34e-02
20	116	4.2	426	1 W08921	Mouse interleukin-12 r	1.38e+00
21	114	4.1	551	1 R06645	Human interleukin-2 re	1.95e+00
22	115	4.1	576	1 R78613	Expression vector pME1	1.64e+00
23	115	4.1	596	1 R78616	Expression vector pME1	1.64e+00

24	115	4.1	600	1 R78610	Expression vector pME1	1.64e+00
25	115	4.1	600	1 R25256	Fas antigen #1.	1.64e+00
26	115	4.1	878	1 R78608	Murine IL-3 receptor b	1.64e+00
27	113	4.1	897	1 R20982	Sequence of beta chain	2.31e+00
28	111	4.0	237	1 R22229	Truncated human growth	3.25e+00
29	111	4.0	269	1 R05045	Soluble human growth h	3.25e+00
30	112	4.0	551	1 R07506	IL-2R beta chain.	2.74e+00
31	112	4.0	592	1 R25227	Fas antigen #2.	2.74e+00
32	110	3.9	237	1 R24274	Truncated human growth	3.85e+00
33	110	3.9	551	1 W39210	Human interleukin-2 re	3.85e+00
34	110	3.9	637	1 P92108	Human growth hormone r	3.85e+00
35	110	3.9	638	1 W33394	Human growth hormone r	3.85e+00
36	109	3.9	878	1 R92529	Fas sequence from AIC2	4.56e+00
37	107	3.8	426	1 W09822	Human interleukin-12 r	6.38e+00
38	107	3.8	427	1 W24973	Human interleukin-13 a	6.38e+00
39	107	3.8	459	1 R08329	Human IL-7 receptor cl	6.38e+00
40	105	3.8	572	1 Y04954	Mycobacterium species	8.91e+00
41	106	3.8	1073	1 W37371	Human ST receptor prot	7.54e+00
42	106	3.8	1073	1 W32063	Human Zcytor2 cytokine	7.54e+00
43	102	3.7	380	1 W36613	Human growth hormone r	1.24e+01
44	103	3.7	638	1 P81326	Armenian hamster alpha	1.05e+01
45	104	3.7	1183	1 Y07728		

ALIGNMENTS

RESULT 1  
ID W31646 standard; Protein; 482 AA.

AC	W31646;				
DT	21-MAY-1998	(first entry)			
DE	Human cytokine receptor gc chain-Ig fusion protein.				
KW	Cytokine receptor; gamma common chain; gc chain; human;				
KW	blocking agent; monoclonal antibody; CpB8; immunological disease;				
KW	myasthenia gravis; rheumatoid arthritis; lupus; multiple sclerosis;				
KW	insulin-dependent diabetes; inflammatory bowel disease;				
KW	sympathetic ophthalmia; uveitis; allergy; asthma; infection;				
KW	graft versus host disease; psoriasis; immunosuppressive; therapy.				
OS	Chimeric - Homo sapiens.				
EH	Key Location/Qualifiers				
FT	Protein	1..254			
FT		/note="gc chain N-terminal region"			
FT	Protein	255..482			
FT		/note="IgG1 constant region"			
FT	Region	255..264			
FT		/note="IgG1 hinge region"			
FT	Domain	264..482			
FT		/note="IgG1 CH2 and CH3 constant domains1"			
PN	WO9743416-A1.				
PD	20-NOV-1997.				
PF	09-MAY-1997; U07870.				
PR	10-MAY-1996; US-017466.				
PA	(BIOJ ) BIOGEN INC.				
PI	Benjamin CD, Burkly LC, Hession C, Whitty A;				
DR	WPI: 98-008885/01.				
DR	N-PSDB; T97439.				
PT	Blocking agents of the gamma common chain of cytokine receptors -				
PT	particularly monoclonal antibodies, used to induce T cell anergy for				
PT	treatment of immunological diseases				
PS	Example 1: Page 79-80; 11pp; English.				
CC	This polypeptide comprises a fusion between the N-terminal 254				
CC	amino acids of the human mature cytokine receptor gamma common (gc)				
CC	chain and the hinge region and CH2 and CH3 constant domains of				
CC	human IgG1. The fusion was expressed from clone pL8001 (see				
CC	T97439) in COS-7 cells, and used to generate murine anti-human gc				
CC	specific monoclonal antibodies (MABs), including CP.B8 produced by				
CC	hybridoma ATCC HB 12107. The invention provides compositions and				
CC	methods for inhibiting cytokine signalling using gc chain blocking				
CC	agents for the treatment of immunological diseases such as				
CC	myasthenia gravis, rheumatoid arthritis, lupus, multiple sclerosis,				
CC	insulin-dependent diabetes, inflammatory bowel disease, syphatic				
CC	ophthalmia, uveitis, allergy, asthma, parasitic infection, graft				
CC	vs. host disease or psoriasis. A preferred gc blocking agent is				
CC	MAB CP.B8 or its Fab fragment (see also W31647-48).				

SQ Sequence 482 AA;  
 Query Match 7.2%; Score 200; DB 1; Length 482;  
 Best Local Similarity 25.7%; Pred. No. 2.25e-07;  
 Matches 53; Conservative 53; Mismatches 88; Indels 12; Gaps 12;  
 Db 59 EVQCFFVNEVMNCTWNSSEPOPTNLTHLWYKNSDNDKVKCHYLSEITSGCQIQ 118  
 QY 31 QIQIYFNLETVQVTWNAKYSR-TNLTFFHYRF-NGD-EAYDOCTNLLQEGHTSGCLLD 87  
 Db 119 KKEHLXQTFVVLQDPPREPRQATQMLKQNLVLPWAPENLTLHLKLSOLELNNRRF 178  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYIL-KPSSPKHVRP-SWHQDAVTVTCSDLS 142  
 Db 179 LNHCLHLVQYRTDWDHSWTEQSDYRHKFSLPSVDGQKRYFRVRSR-FNPLCGSAQHW 237  
 QY 143 YGD-LLYEVOYRSPFDTEW-QSKOENTCNVTIEGLDAEKCYSFVVRVYKAMEDVYGPDTYP 200  
 Db 238 SEWSHPITHWGSNTSKENVDKTHTCPP 263  
 QY 201 SDWSEVTCWQGEIRDAETPT-PP 225

RESULT 2  
 ID R82934 standard; Protein; 230 AA.  
 AC R82934;  
 DT 26-FEB-1996 (first entry)  
 DE Interleukin-4 component common to the IL-2 receptor gamma chain.  
 KW Interleukin-4; IL-4; gamma chain component; immunosuppressants;  
 KW anti-allergy agent; signal transduction inhibitor; autoimmune;  
 KW disease; anti-inflammatory; anaphylactic shock; bronchial asthma;  
 KW Interleukin-2; IL-2; atopic dermatitis; urticaria.  
 OS Homo sapiens.  
 PN J07149662-A.  
 PD 13-JUN-1995.  
 PR 07-SEP-1994; 213706.  
 PR 08-SEP-1993; JP-223574.  
 PA (AJIN) AJINOMOTO KK.  
 PA (SUGA) SUGAMURA K.  
 DR WPI: 95-243601/32.  
 DR N-PSDB: T04952.  
 PT Novel interleukin-4 receptor monoclonal antibodies inhibit signal  
 PT transmission - useful as immunosuppressants and anti-allergy agents.  
 PS Example 1; Page 9; ilpp; Japanese.  
 CC T04952 encodes R82934 a component of the IL-4 receptor common to  
 CC the IL-2 receptor gamma chain molecule, which was used to generate  
 CC anti-IL-4 receptor monoclonal antibodies (mAbs). The mAbs (IL-4  
 CC signal transduction inhibitors) can be used as immunosuppressants  
 CC and anti-allergy agents, for the treatment of autoimmune and chronic  
 CC inflammatory diseases, e.g. anaphylactic shock, bronchial asthma,  
 CC atopic dermatitis and urticaria.  
 SQ Sequence 230 AA;

Query Match 7.0%; Score 195; DB 1; Length 230;  
 Best Local Similarity 25.9%; Pred. No. 5.99e-07;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 38 EVQCFFVNEVMNCTWNSSEPOPTNLTHLWYKNSDNDKVKCHYLSEITSGCQIQ 97  
 QY 31 QIQIYFNLETVQVTWNAKYSR-TNLTFFHYRF-NGD-EAYDOCTNLLQEGHTSGCLLD 87  
 Db 98 KKEHLXQTFVVLQDPPREPRQATQMLKQNLVLPWAPENLTLHLKLSOLELNNRRF 157  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYIL-KPSSPKHVRP-SWHQDAVTVTCSDLS 142  
 Db 158 LNHCLHLVQYRTDWDHSWTEQSDYRHKFSLPSVDGQKRYFRVRSR-FNPLCGSAQHW 216  
 QY 143 YGD-LLYEVOYRSPFDTEW-QSKOENTCNVTIEGLDAEKCYSFVVRVYKAMEDVYGPDTYP 200  
 Db 217 SEWSHPITHW 225  
 QY 201 SDWSEVTCW 209

RESULT 3  
 ID R47151 standard; Protein; 230 AA.  
 AC R47151;  
 DT 13-JUN-1994 (first entry)  
 DE IL-2 receptor gamma chain.  
 KW Interleukin-2 receptor gamma chain.  
 KW rheumatoid arthritis; transplant rejection; primer;  
 KW polymerase chain reaction; PCR; amplification.  
 OS Homo sapiens.  
 PN EP-578932-A.  
 PD 19-JAN-1994.  
 PR 22-APR-1993; 106561.  
 PR 23-APR-1992; JP-104947.  
 PA (AJIN) AJINOMOTO KK.  
 PA (SUGA) SUGAMURA K.  
 PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
 PI Suzuki M, Takeshita T;  
 DR WPI: 94-017546/03.  
 DR P-PSDB: Q54831.  
 PT DNA and protein sequences of IL-2 gamma chain - useful as immune  
 PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
 PT transplant rejection  
 PS Disclosure; Page 22-23, 35-36; 50pp; English.  
 CC The human IL-2 receptor gamma chain preform (R47148), including the  
 CC signal peptide, is encoded by the sequence given in Q54828. The  
 CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
 CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
 CC while a soluble form suitable for expression in prokaryotes (R47151)  
 CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
 CC terminal sequence of IL-2 receptor gamma chain, and are used to  
 CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
 CC are used to obtain the protein given in R47151.  
 SQ Sequence 230 AA;

Query Match 7.0%; Score 195; DB 1; Length 230;  
 Best Local Similarity 25.9%; Pred. No. 5.99e-07;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 37 EVQCFFVNEVMNCTWNSSEPOPTNLTHLWYKNSDNDKVKCHYLSEITSGCQIQ 96  
 QY 31 QIQIYFNLETVQVTWNAKYSR-TNLTFFHYRF-NGD-EAYDOCTNLLQEGHTSGCLLD 87  
 Db 97 KKEHLXQTFVVLQDPPREPRQATQMLKQNLVLPWAPENLTLHLKLSOLELNNRRF 156  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYIL-KPSSPKHVRP-SWHQDAVTVTCSDLS 142  
 Db 157 LNHCLHLVQYRTDWDHSWTEQSDYRHKFSLPSVDGQKRYFRVRSR-FNPLCGSAQHW 215  
 QY 143 YGD-LLYEVOYRSPFDTEW-QSKOENTCNVTIEGLDAEKCYSFVVRVYKAMEDVYGPDTYP 200  
 Db 216 SEWSHPITHW 224  
 QY 201 SDWSEVTCW 209  
 RESULT 4  
 ID R47150 standard; Protein; 252 AA.  
 AC R47150;  
 DT 13-JUN-1994 (first entry)  
 DE IL-2 receptor gamma chain.  
 KW Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;  
 KW rheumatoid arthritis; transplant rejection; primer;  
 KW polymerase chain reaction; PCR; amplification; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Peptide 1..22 /label= sig\_peptide  
 PN EP-578932-A.  
 PD 19-JAN-1994.  
 PR 22-APR-1993; 106561.  
 PR 23-APR-1992; JP-104947.  
 PA (AJIN) AJINOMOTO KK.

PA (SUGA/) SUGAMURA K.  
PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
DR Suzuki M, Takeshita T;  
DR WPI: 94-017546/03.  
DR P-PSDB: Q54830.  
PT DNA and protein sequences of IL-2 gamma chain - useful as immune  
PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
PT transplant rejection  
PS Disclosure: Page 21-22, 34-35; 50pp; English.  
CC The human IL-2 receptor gamma chain preform (R47148), including the  
CC signal peptide, is encoded by the sequence given in Q54828. The  
CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
CC while a soluble form suitable for expression in prokaryotes (R47151)  
CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
CC terminal sequence of IL-2 receptor gamma chain, and are used to  
CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
CC are used to obtain the protein given in R47151.  
SQ Sequence 252 AA;  
Query Match 7.0%; Score 195; DB 1; Length 252;  
Best Local Similarity 25.9%; Pred. No. 5.99e-07;  
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
Db 59 EVQCFVENVYMNCTWNSSEPPQNTLTHYWKNSDNDKVKCSHYLFSSEITSGCOLQ 118  
QY 31 QIQIYFNLETQVTVWNASKYSR-TNLTFHYRF-NGD-EAYDOCTNYLLQEGHTSGCLLD 87  
Db 119 KKEIHLVQTFVVOQLDPPRPRQATOMLKLQNLVWPAPENLTLHLKLSOLELNWNNRF 178  
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYIL-KPSSPKHVRF-SWHQDAVITVCSCLS 142  
Db 179 LNHCLHLVQYRTDWDHSHWTEQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 237  
QY 143 YGD-LLYEVQYRSPDTEW-OSKQENTCNVTIEGLDAEKCSFWVRKAMEDVYGPDTYP 200  
Db 238 SEWSHPHFW 246  
QY 201 SDWSEVTCW 209  
RESULT 5  
ID R47149 standard; Protein; 347 AA.  
AC R47149 (first entry)  
DE IL-2 receptor gamma chain.  
KW Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;  
KW rheumatoid arthritis; transplant rejection; primer;  
KW polymerase chain reaction; PCR; amplification.  
OS Homo sapiens.  
PN EP-578932-A.  
PD 19-JAN-1994.  
PF 22-APR-1993; 106561.  
PR 23-APR-1992; JP-104947.  
PA (AJIN) AJINOMOTO KK.  
PA (SUGA/) SUGAMURA K.  
PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
PI Suzuki M, Takeshita T;  
PI WPI: 94-017546/03.  
PI P-PSDB: Q54828.  
PT DNA and protein sequences of IL-2 gamma chain - useful as immune  
PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
PT transplant rejection  
PS Disclosure: Page 41; 50pp; English.  
CC The human IL-2 receptor gamma chain preform (R47148), including the  
CC signal peptide, is encoded by the sequence given in Q54828. The  
CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
CC while a soluble form suitable for expression in prokaryotes (R47151)  
CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
CC terminal sequence of IL-2 receptor gamma chain, and are used to  
CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
CC are used to obtain the protein given in R47151.  
SQ Sequence 347 AA;  
Query Match 7.0%; Score 195; DB 1; Length 347;  
Best Local Similarity 25.9%; Pred. No. 5.99e-07;  
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
Db 59 EVQCFVENVYMNCTWNSSEPPQNTLTHYWKNSDNDKVKCSHYLFSSEITSGCOLQ 96  
QY 31 QIQIYFNLETQVTVWNASKYSR-TNLTFHYRF-NGD-EAYDOCTNYLLQEGHTSGCLLD 87  
Db 97 KKEIHLVQTFVVOQLDPPRPRQATOMLKLQNLVWPAPENLTLHLKLSOLELNWNNRF 156  
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYIL-KPSSPKHVRF-SWHQDAVITVCSCLS 142  
Db 157 LNHCLHLVQYRTDWDHSHWTEQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 215  
QY 143 YGD-LLYEVQYRSPDTEW-OSKQENTCNVTIEGLDAEKCSFWVRKAMEDVYGPDTYP 200  
Db 216 SEWSHPHFW 224  
QY 201 SDWSEVTCW 209

SQ Sequence 347 AA;  
Query Match 7.0%; Score 195; DB 1; Length 347;  
Best Local Similarity 25.9%; Pred. No. 5.99e-07;  
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
Db 37 EVQCFVENVYMNCTWNSSEPPQNTLTHYWKNSDNDKVKCSHYLFSSEITSGCOLQ 96  
QY 31 QIQIYFNLETQVTVWNASKYSR-TNLTFHYRF-NGD-EAYDOCTNYLLQEGHTSGCLLD 87  
Db 97 KKEIHLVQTFVVOQLDPPRPRQATOMLKLQNLVWPAPENLTLHLKLSOLELNWNNRF 156  
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYIL-KPSSPKHVRF-SWHQDAVITVCSCLS 142  
Db 157 LNHCLHLVQYRTDWDHSHWTEQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 215  
QY 143 YGD-LLYEVQYRSPDTEW-OSKQENTCNVTIEGLDAEKCSFWVRKAMEDVYGPDTYP 200  
Db 216 SEWSHPHFW 224  
QY 201 SDWSEVTCW 209  
RESULT 6  
ID R47148 standard; Protein; 369 AA.  
AC R47148 (first entry)  
DE IL-2 receptor gamma chain.  
KW Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;  
KW rheumatoid arthritis; transplant rejection; primer;  
KW polymerase chain reaction; amplification; ss.  
OS Homo sapiens.  
PN EP-578932-A.  
PD 19-JAN-1994.  
PF 22-APR-1993; 106561.  
PR 23-APR-1992; JP-104947.  
PA (AJIN) AJINOMOTO KK.  
PA (SUGA/) SUGAMURA K.  
PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
PI Suzuki M, Takeshita T;  
PI WPI: 94-017546/03.  
PI P-PSDB: Q54828.  
PT DNA and protein sequences of IL-2 gamma chain - useful as immune  
PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
PT transplant rejection  
PS Disclosure: Page 16-17, 29-30; 50pp; English.  
CC The human IL-2 receptor gamma chain preform (R47148), including the  
CC signal peptide, is encoded by the sequence given in Q54828. The  
CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
CC while a soluble form suitable for expression in prokaryotes (R47151)  
CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
CC terminal sequence of IL-2 receptor gamma chain, and are used to  
CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
CC are used to obtain the protein given in R47151.  
SQ Sequence 369 AA;  
Query Match 7.0%; Score 195; DB 1; Length 369;  
Best Local Similarity 25.9%; Pred. No. 5.99e-07;  
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
Db 59 EVQCFVENVYMNCTWNSSEPPQNTLTHYWKNSDNDKVKCSHYLFSSEITSGCOLQ 118  
QY 31 QIQIYFNLETQVTVWNASKYSR-TNLTFHYRF-NGD-EAYDOCTNYLLQEGHTSGCLLD 87  
Db 119 KKEIHLVQTFVVOQLDPPRPRQATOMLKLQNLVWPAPENLTLHLKLSOLELNWNNRF 178  
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYIL-KPSSPKHVRF-SWHQDAVITVCSCLS 142  
Db 179 LNHCLHLVQYRTDWDHSHWTEQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 237

QY 143 YGD-LLVEYQYRSPDTEW-QSKQENTCNVTIEGLDAEKCYFWRVRKAMEDVIGPDTYP 200  
 Db 238 SEWSHPHWH 246  
 QY 201 SDWSEVTCW 209

RESULT 7  
 ID R59094 standard; Protein; 369 AA.  
 AC R59094;  
 DT 04-MAY-1995 (first entry)  
 DE Murine IL-2R gamma.  
 KW Murine IL-2R gamma: X-linked severe combined immunodeficiency;  
 KW XSCID; interleukin.  
 OS Mus musculus.  
 FH Key

FT peptide  
 FT 1. .21 Location/Qualifiers  
 FT domain  
 FT 258. .284 /note= "signal peptide"  
 FT misc\_difference 331 /note= "transmembrane domain"  
 FT modified\_site 71. .73 /note= "Corresponding codon CAG"  
 FT modified\_site 75. .77 /label= N-glycosylation\_site  
 FT modified\_site 84. .86 /label= N-glycosylation\_site  
 FT modified\_site 96. .98 /label= N-glycosylation\_site  
 FT modified\_site 159. .161 /label= N-glycosylation\_site  
 FT modified\_site 255. .257 /label= N-glycosylation\_site  
 FT modified\_site 255. .257 /label= N-glycosylation\_site  
 FT W09420641-A.  
 PD 15-SEP-1994.  
 PR 10-MAR-1994; U02891.  
 PR 12-MAR-1993; US-031143.  
 PR 14-SEP-1993; US-121435.  
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 PI Leonard WJ, McBride WO, Noguchi M;  
 DR WPI; 94-303046/37.  
 DR N-PSDB; Q71977.

PT Diagnosis of X-linked severe combined immunodeficiency (XSCID) -  
 PT comprises detecting mutated IL-2R gamma gene, also vectors and  
 PT transgenic animals containing the mutated gene  
 PS Example 1; Fig 7; 98pp; English.  
 CC Q71977 is the DNA sequence of murine IL-2R gamma R59094,  
 CC this was used in the development of a claimed method for the  
 CC diagnosis of X-linked severe combined immunodeficiency (XSCID),  
 CC in female carriers and male sufferers.  
 SQ Sequence 369 AA;

Query Match 6.7%; Score 188; DB 1; Length 369;  
 Best Local Similarity 26.8%; Pred. No. 2,34e-06;  
 Matches 51; Conservative 52; Mismatches 75; Indels 12; Gaps 11;

Db 59 EVQCFVNEVYMNCTSSSPQATNLTLHYRYKVSDDNTFQECSHYLFSEIKSGCOIQ 118  
 QY 31 QIQIIFYNLEVQVWNAKYSR-TNLTFFHFRN-GDE-AYDOCTNYLLQEGHTSGCLLD 87  
 Db 119 KEDIQYQTFVVLQDQPKORRAVQKLNQNLVPRAPENLTLSNLSESOLELRWKSRR 178  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-NVYL-KPSPKHYRFS-WHQDQVTV--TCSD 140  
 Db 179 IKERCLOYLVYRNRDRSWTELIVNHEPRELSVLDLKYTFYRSRY-NPICGSSQ 237  
 QY 141 LSYGDLLEYQYRSPDTEW-QSKQENTCNVTIEGLDAEKCYFWRVRKAMEDVIGPDTY 199  
 Db 238 WSKWSQPVH 247  
 QY 200 PSDWSEVTCW 209

RESULT 8  
 ID W64057 standard; Protein; 500 AA.  
 AC W64057;  
 DT 02-OCT-1998 (first entry)  
 DE Human IL-9 receptor protein variant #2.  
 KW Interleukin-9 receptor; IL-9R; Asthma associated factor 2; AAF2;  
 KW soluble; bronchial hyper-responsiveness; detection; diagnosis; human;  
 KW eosinophil; HL60 cell; antagonist.  
 OS Homo sapiens.  
 PN W09824904-A2.  
 PD 11-JUN-1998.  
 PF 02-DEC-1997; U21992.  
 PR 01-DEC-1997; US-032224.  
 PR 02-DEC-1996; US-032224.  
 PA (MAGA-) MAGAININ PHARM INC.  
 PI Grasso L, Holroyd KJ, Levitt RC, Nicolaides NC;  
 DR WPI; 98-348150/30.  
 PT Human interleukin-9 receptor variant(s) useful in treating asthma -  
 PT and related disorders e.g. bronchial hyper-responsiveness and for  
 PT diagnosing greater or alternatively less susceptibility to these  
 PT conditions  
 PS Disclosure; Page -; 81pp; English.  
 CC This sequence represents a human interleukin-9 (IL-9) receptor variant  
 CC (also known as Asthma Associated Factor 2 or AAF2) which is used in the  
 CC construction of protein variants which are soluble and can be  
 CC administered to humans to alleviate asthma and related disorders e.g.  
 CC bronchial hyper-responsiveness. The DNA molecules are useful to detect  
 CC or diagnose susceptibility to such conditions. Cells used in this method  
 CC may be e.g. eosinophils and HL60 cells. Cells lacking human IL-9 receptor  
 CC can be transfected with the DNA and used to identify IL-9 pathway  
 CC antagonists. The nucleic acids can also be used to design probes to  
 CC detect other IL-9 receptor variants.  
 SQ Sequence 500 AA;

Query Match 6.5%; Score 182; DB 1; Length 500;  
 Best Local Similarity 26.9%; Pred. No. 7.45e-06;  
 Matches 61; Conservative 55; Mismatches 87; Indels 24; Gaps 20;  
 Db 157 LLSYELAFKKQ-EEAEWAQHRDHVGVWTLLEAFELDPGFIHEARLVQVATLEDDVVE 215  
 QY 146 LL-VEVQYRSPDTEWQSKQ-ENTCNVT--I-EGLDAEKCYSF--WVRV-KA-MED-VYG 195  
 Db 216 EERTQWSE---WSQPVCFQAPQORQPLPPWGPNG-TLVAVSIFLLLTGPTVL-LFK 270  
 QY 196 PDTYPSDWSEVTCQGE-IRDAEATP-TTPKPKLSKFISSIAILLWSLLLSLWK 253  
 Db 271 LSPRVKRFYQNVSPAMFFQPLYSVHNGFQTMGAHRAGVLLSQDCAGTPQGALEPCV 330  
 QY 254 LW-RYKFKFLIPSPDPKSIFFGLFEIHQGNFQEWITDTQNVHL-HKMAGAEQES-GPEE 310  
 Db 331 QEATALLTCGPAPRWKPSVAAEEQGPQ-T-RLPGN-LSSSEDLVLPAG 374  
 QY 311 PLVQLAKTEASPRMLDPQTEERASGSLQPLPQLOGDGVVTIG 357

RESULT 9  
 ID W64055 standard; Protein; 501 AA.  
 AC W64055;  
 DT 02-OCT-1998 (first entry)  
 DE Human IL-9 receptor protein.  
 KW Interleukin-9 receptor; IL-9R; Asthma associated factor 2; AAF2;  
 KW soluble; bronchial hyper-responsiveness; detection; diagnosis; human;  
 KW eosinophil; HL60 cell; antagonist.  
 OS Homo sapiens.  
 PN W09824904-A2.  
 PD 11-JUN-1998.  
 PF 02-DEC-1997; U21992.  
 PR 01-DEC-1997; US-032224.  
 PR 02-DEC-1996; US-032224.  
 PA (MAGA-) MAGAININ PHARM INC.  
 PI Grasso L, Holroyd KJ, Levitt RC, Nicolaides NC;

DR WPI: 98-348150/30.  
DR N-PSDB; V44088.  
PT Human interleukin-9 receptor variant(s) useful in treating asthma -  
PT and related disorders e.g. bronchial hyper-responsiveness and for  
PT diagnosing greater or alternatively less susceptibility to these  
PT conditions. Fig 2: 81pp; English.  
PT Disclosure: represents a human interleukin-9 (IL-9) receptor (also  
CC known as Asthma Associated Factor 2 or AAF2) which is used in the  
CC construction of protein variants which are soluble and can be  
CC administered to humans to alleviate asthma and related disorders e.g.  
CC bronchial hyper-responsiveness. The DNA molecules are useful to detect  
CC or diagnose susceptibility to such conditions. Cells used in this method  
CC may be e.g. eosinophils and H160 cells. Cells lacking human IL-9 receptor  
CC can be transfected with the DNA and used to identify IL-9 pathway  
CC antagonists. The nucleic acids can also be used to design probes to  
CC detect other IL-9 receptor variants.  
CC Sequence 501 AA;  
SQ

Query Match 6.3%; Score 176; DB 1; Length 501;  
Best Local Similarity 27.6%; Pred. No. 2.36e-05;  
Matches 50; Conservative 43; Mismatches 74; Indels 14; Gaps 12;

Db 203 LRQMATLEDDVVEERYTGOWSE---WSQPVCFQAPQROGLPPGWPGN-TLVAVSI 258  
QY 183 FWRVVKAMED-VYGPDTYPSDWSEVTCWQGE-IRDAEAETP-TTPKPKLSKFLIISLA 239  
Db 259 FLITGTPTYL-LFKLSPRVKRIEYQNVPSAMFFQPLYSVHNGNFQTMGAHRAVLSQ 317  
QY 240 ILLVSVLLLLSLWKLW-RVKKFLIPSPDKSIFFGLFEIHQGNFQEWITDTQNVHL-H 297  
Db 318 DCAGTPQALPCVQEQATALLTCGPAPRWKSVALEEEOEGPG-T-RLPGN-LSSSEDLVPA 374  
QY 298 KMGAEQES-GPEPLVVLQAKTEAESPRLDPQTEKEASGSLQLPHQPLQGGDVVTI 356  
Db 375 G 375  
QY 357 G 357

RESULT 10  
ID W64056 standard; Protein; 501 AA.  
AC W64056;  
DT 02-OCT-1998 (first entry)  
DE Human IL-9 receptor protein variant #1.  
KW Interleukin-9 receptor; IL-9R; Asthma associated factor 2; AAF2;  
KW soluble; bronchial hyper-responsiveness; detection; diagnosis; human;  
KW eosinophil; H160 cell; antagonist.  
OS Homo sapiens.  
PN W098249042.  
PD 11-JUN-1998; U21992.  
PF 02-DEC-1997; U21992.  
PR 01-DEC-1997; US-032224.  
PR 02-DEC-1996; US-032224.  
PA (MAGS) MAGANIN PHARM INC.  
PI Grasso L, Holroyd KJ, Levitt RC, Nicolaides NC;  
DR WPI: 98-348150/30.  
PT Human interleukin-9 receptor variant(s) useful in treating asthma -  
PT and related disorders e.g. bronchial hyper-responsiveness and for  
PT diagnosing greater or alternatively less susceptibility to these  
PT conditions. Page -; 81pp; English.  
PT Disclosure: represents a human interleukin-9 (IL-9) receptor variant  
CC (also known as Asthma Associated Factor 2 or AAF2) which is used in the  
CC construction of protein variants which are soluble and can be  
CC administered to humans to alleviate asthma and related disorders e.g.  
CC bronchial hyper-responsiveness. The DNA molecules are useful to detect  
CC or diagnose susceptibility to such conditions. Cells used in this method  
CC may be e.g. eosinophils and H160 cells. Cells lacking human IL-9 receptor  
CC can be transfected with the DNA and used to identify IL-9 pathway  
CC antagonists. The nucleic acids can also be used to design probes to  
CC detect other IL-9 receptor variants.  
CC Sequence 501 AA;  
SQ

Query Match 6.3%; Score 175; DB 1; Length 501;  
Best Local Similarity 27.6%; Pred. No. 2.86e-05;  
Matches 50; Conservative 43; Mismatches 74; Indels 14; Gaps 12;

Db 203 LRQMATLEDDVVEERYTGOWSE---WSQPVCFQAPQROGLPPGWPGN-TLVAVSI 258  
QY 183 FWRVVKAMED-VYGPDTYPSDWSEVTCWQGE-IRDAEAETP-TTPKPKLSKFLIISLA 239  
Db 259 FLITGTPTYL-LFKLSPRVKRIEYQNVPSAMFFQPLYSVHNGNFQTMGAHRAVLSQ 317  
QY 240 ILLVSVLLLLSLWKLW-RVKKFLIPSPDKSIFFGLFEIHQGNFQEWITDTQNVHL-H 297  
Db 318 DCAGTPQALPCVQEQATALLTCGPAPRWKSVALEEEOEGPG-T-RLPGN-LSSSEDLVPA 374  
QY 298 KMGAEQES-GPEPLVVLQAKTEAESPRLDPQTEKEASGSLQLPHQPLQGGDVVTI 356  
Db 375 G 375  
QY 357 G 357

RESULT 11  
ID R47517 standard; Protein; 507 AA.  
AC R47517;  
DT 24-JUN-1994 (first entry)  
DE MEL EPO receptor.  
KW Erythropoietin receptor; recombinant; murine; anaemia.  
OS Mus musculus.  
PH Key Location/Qualifiers  
FT peptide 1..24 /note= "signal"  
FT protein 25..507 /note= "mature EPO receptor"  
FT modified\_site 75 /note= "potential N-glycosylation site"  
FT modified\_site 383 /note= "potential N-glycosylation site"  
FT region 250..271 /note= "putative transmembrane region"  
PN US5278065-A.  
PD 11-JAN-1994; 306503.  
PF 03-FEB-1989; US-306503.  
PR 25-MAR-1991; US-678877.  
PA (CHIL-) CHILDRENS MEDICAL CENT.  
PA (GEM) GENETICS INST INC.  
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.  
PI D'andrea A, Jones SS, Wong GG;  
DR WPI: 94-025409/03.  
DR N-PSDB; Q53994.  
PT Recombinant DNA encoding erythropoietin receptor - used to  
PT develop prods. for study, treatment or diagnosis of disorders in  
PT which receptor is dysfunctional  
PS Disclosure: Fig 2: 24pp; English.  
CC Mouse erythroleukaemia (MEL) cells were used to construct a cDNA  
CC library. The cDNA was used to transfect COS-1 cells and these were  
CC screened for radioiodinated erythropoietin (EPO) binding to isolate  
CC cDNA encoding the EPO receptor. The cDNA may be used to isolate the  
CC EPO receptor from other sources and to study, treat or diagnose the  
CC disorders in which the EPO receptor is dysfunctional. The EPO  
CC receptor may also be used to raise antibodies for treating  
CC hypersensitivity to EPO or who have elevated levels of EPO. The prod.  
CC is pref. used for treating anaemias, primary proliferative polycythemia  
CC and secondary polycythemia.  
CC See also R47518.  
SQ Sequence 507 AA;  
Query Match 5.8%; Score 163; DB 1; Length 507;  
Best Local Similarity 46.8%; Pred. No. 2.79e-04;  
Matches 29; Conservative 9; Mismatches 20; Indels 4; Gaps 3;

Db 247 LDPLILTLISL-ILVILSLTLVALLSHRRTLOOKIWPGIPSPSEFEGLTTHKGNFQL 305

QY 229 LSKFLLSSAILLMVSL--LSLWKLVRVKFLI-PSVPDPKSIFFGLFEIHQGNFQE 285  
 Db 306 WL 307  
 QY 286 WI 287

RESULT 12  
 ID R69502 standard; Protein; 507 AA.

AC R69502;  
 DT 10-AUG-1995 (first entry)  
 DE Mouse erythropoietin receptor.  
 KW Erythropoietin receptor; anemia therapy; signal peptide;  
 KW transmembrane region; N-linked glycosylation.  
 OS Mus musculus.  
 FH Key  
 FT peptide 1..24  
 FT protein /note= "signal peptide"  
 FT protein 25..507  
 FT modified\_site /note= "mature protein"  
 FT modified\_site 75..77  
 FT domain /note= "N-linked glycosylation site"  
 FT domain 250..271  
 FT modified\_site /note= "transmembrane region"  
 FT modified\_site 383..385  
 FT /note= "N-linked glycosylation site"  
 PN US5378808-A.  
 PD 03-JAN-1995.  
 PF 03-FEB-1989; 306503.  
 PR 03-FEB-1989; US-306503.  
 PR 25-MAR-1991; US-678877.  
 PR 10-JUN-1993; US-075069.  
 PA (GENE ) GENETICS INST INC.  
 PI D'andrea A, Jones SS, Wong GG;  
 DR WPI: 95-051310/07.  
 DR N-PSDB; Q81891.  
 PT New recombinant erythropoietin receptor polypeptide(s) - used for  
 PT detection, purification, and therapy and for prodn. of antibodies for  
 PT detection and therapy  
 PS Claim 1; Fig 2: 24pp; English.  
 CC The sequence corresponds to a mouse erythropoietin receptor,  
 CC including putative signal peptide and transmembrane regions, and 2  
 CC N-linked glycosylation sites. The protein is derived from mouse  
 CC erythroleukemia cells and may be used in drug design or in  
 CC pharmaceutical compositions for therapy of anemia.  
 SQ Sequence 507 AA;

Query Match 5.8%; Score 163; DB 1; Length 507;  
 Best Local Similarity 46.8%; Pred. No. 2.79e-04;  
 Matches 29; Conservative 9; Mismatches 20; Indels 4; Gaps 3;

Db 247 LDPLILTL--LIVLISLLLTVALLSHRTLOQKIWPICSPSEFEGLTTHKGNFQL 305  
 QY 229 LSKFLLSSAILLMVSL--LSLWKLVRVKFLI-PSVPDPKSIFFGLFEIHQGNFQE 285  
 Db 306 WL 307  
 QY 286 WI 287

RESULT 13  
 ID R50327 standard; Protein; 507 AA.

AC R50327;  
 DT 19-OCT-1994 (first entry)  
 DE Mouse soluble EPO receptor protein.  
 KW Murine; soluble; erythropoietin; EPO; receptor protein; SEPO-R; drug;  
 KW antigen; diagnostic agent; biochemical reagent.  
 OS Mus musculus.  
 FH Key  
 FT modified\_site 75..77  
 FT /note= "N-linked glycosylation site"  
 PN J06038787-A.

PD 15-FEB-1994. 082865.  
 PF 04-MAR-1992; JP-082865.  
 PR 04-MAR-1992; JP-082865.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 DR WPI: 94-094847/12.  
 DR N-PSDB; Q44854.  
 PT Soluble erythropoietin receptor protein - and DNA coding for  
 PT SEPO-R, useful as diagnostic reagent  
 PS Disclosure; Fig 1; 9pp; Japanese.  
 CC This sequence represents the murine soluble erythropoietin (EPO)  
 CC receptor protein (SEPO-R). This protein is able to bind to EPO and  
 CC has antigenicity as an EPO receptor. The molecular weight of the  
 CC full length protein is pref 33 or 29 kD. The protein is useful as a  
 CC drug, as a diagnostic agent and a biochemical reagent.  
 SQ Sequence 507 AA;

Query Match 5.8%; Score 163; DB 1; Length 507;  
 Best Local Similarity 46.8%; Pred. No. 2.79e-04;  
 Matches 29; Conservative 9; Mismatches 20; Indels 4; Gaps 3;

Db 247 LDPLILTL--LIVLISLLLTVALLSHRTLOQKIWPICSPSEFEGLTTHKGNFQL 305  
 QY 229 LSKFLLSSAILLMVSL--LSLWKLVRVKFLI-PSVPDPKSIFFGLFEIHQGNFQE 285  
 Db 306 WL 307  
 QY 286 WI 287

RESULT 14

ID R06511 standard; protein; 507 AA.  
 AC R06511;  
 DT 04-JAN-1991 (first entry)  
 DE EPO receptor sequence deduced from DNA of clone 190.  
 KW Erythropoietin; Diamond Blackfan anaemia; polycythemia vera.  
 OS Mus musculus.

FH Key  
 FT peptide 1..24  
 FT domain /label=signal peptide  
 FT domain 25..248  
 FT /label=extracellular domain  
 FT domain /note=EPO binding region  
 FT domain 248..271  
 FT domain /label=transmembrane domain  
 FT domain 272..507  
 FT modified\_site /label=intracellular domain  
 FT modified\_site 75..77  
 FT /label=N-linked-glycos  
 FT modified\_site 182..184  
 FT /label=N-linked-glycos

PN W09008822-A.

PD 09-AUG-1990.  
 PF 01-FEB-1990; U00635.  
 PR 03-FEB-1989; US-306503.  
 PA (GENE-) GENETICS INST INC.  
 PA (WHIT-) WHITEHEAD INST.  
 PI D'andrea A, Wong G;  
 DR WPI: 90-260931/34.  
 DR N-PSDB; Q05747.

PT Erythropoietin receptor and gene - used for developing reagents  
 PT and systems to control and study erythropoiesis.  
 PS Disclosure; Fig 1; 53pp; English.  
 CC The sequence was deduced from DNA from a clone isolated from a  
 CC cDNA library prep. from uninduced murine erythroleukemia cells.  
 CC It is a type I transmembrane protein with binding affinity for EPO.  
 CC The gene and recombinant EPO receptor produced on expression of  
 CC the DNA are used to develop reagents and systems to control and  
 CC study erythropoiesis. It is believed that the EPO receptor is  
 CC dysfunctional in individuals with Diamond Blackfan anaemia, and  
 CC may be hyperactive in polycythemia vera.  
 CC See also R06512 (human EPO receptor).  
 SQ Sequence 507 AA;

Thu May 11 06:49:23 2000

Query Match 5.7%; Score 160; DB 1; Length 507;  
Best Local Similarity 46.8%; Pred. No. 4.90e-04;  
Matches 29; Conservative 9; Mismatches 20; Indels 4; Gaps 3;  
Db 247 LDPLILTLISL-ILVILISLLVLLVALLSHRRALQKIWCPSEFEGLETTTHKGNFOL 305  
Qy 229 LSKFILLSSALLLVLSLL--LSLWKLWRVKKFLI-PSVDPKSIFFGLFELHOGNFQE 285  
Db 306 WL 307  
Qy 286 WI 287

RESULT 15  
ID R69503 standard; Protein; 508 AA.  
AC R69503;  
DT 11-AUG-1995 (first entry)  
DE Human erythropoietin receptor;  
KW Erythropoietin receptor; anemia therapy; diagnostic.  
OS Homo sapiens.  
FH key Location/Qualifiers  
FT peptide 1..24 "signal peptide"  
FT protein 23...508  
FT modified\_site 76...79 "mature protein"  
FT domain /note="N-glycosylation site"  
FT /note="transmembrane region"  
PN US5378908-A.  
PD 03-JAN-1995. 306503.  
PF 03-FEB-1989; US-206503.  
PR 03-FEB-1989; US-27887.  
PR 25-MAR-1991; US-075069.  
PR 10-JUN-1993; US-075069.  
PA (GENI) GENETICS INS. INC.  
PI D'Andrea G, Jones SS, Wong GG;  
DR WPI95-04310/07.  
DR N-PSDB; Q81892.  
TI New recombinant erythropoietin receptor polypeptide(s) - used for  
PI detection, purification and therapy and for prodn. of antibodies for  
PI detection and therapy.  
PI Citeam 2; Fig 9, 24pp. English.  
CC The sequence is that of a 55-kDa human erythropoietin receptor. The  
CC receptor polypeptide may be used in purification and detection of  
CC erythropoietin and in production of antibodies for anemia therapy.  
CC the polypeptide may also be used for treating individuals  
CC who are hypersensitive to erythropoietin or who have elevated  
CC erythropoietin levels. They may be used in therapy of e.g. primary  
CC or secondary proliferative polycythemia.  
SQ Sequence 508 AA;

Query Match 5.7%; Score 160; DB 1; Length 508;  
Best Local Similarity 42.6%; Pred. No. 4.90e-04;  
Matches 26; Conservative 12; Mismatches 21; Indels 2; Gaps 2;  
Db 248 LDPLILTLISL-ILVILISLLVLLVALLSHRRALQKIWCPSEFEGLETTTHKGNFOLW 307  
Qy 229 LSKFILLSSALLLVLSLL--LSLWKLWRVKKFLI-PSVDPKSIFFGLFELHOGNFQE 286  
Db 308 L 308  
Qy 287 I 287

Search completed: wed May 10 11:10:19 2000  
Job time : 18 secs.





\*\*\*\*\*  
 M P E R L H  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 14:39:38 2000; MasPar time 12.62 Seconds  
 696.462 Million cell updates/sec  
 Tabular output not generated.

Title: >US-09-376-430-2  
 Description: (1-371) from US09376430A.pap  
 Perfect score: 371  
 Sequence: 1 MGRLVLLGNAVFLGLGWWA.....DVTIGGFTFVNDRSYVAL 371

Scoring table: TABLE unitprotatable  
 Gap 60

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: a-geneseq35  
 1:geneseq

Statistics: Mean 2.917; Variance 0.716; scale 4.075

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	8	2.2	499	1 R94349	Rat cytochrome P450 2C	1.90e+00
2	7	1.9	40	1 W93223	Human cytochrome P450	2.79e+01
3	7	1.9	40	1 W93224	Human cytochrome P450	2.79e+01
4	7	1.9	41	1 W93226	Human cytochrome P450	2.79e+01
5	7	1.9	72	1 Y12020	Human 5' EST secreted	2.79e+01
6	7	1.9	105	1 P50237	Human epidermal growth	2.79e+01
7	7	1.9	112	1 W60650	Human C-C chemokine DG	2.79e+01
8	7	1.9	131	1 W44082	Human secreted protein	2.79e+01
9	7	1.9	131	1 W27646	Secreted protein AK296	2.79e+01
10	7	1.9	393	1 W55091	Streptococcus pneumonia	2.79e+01
11	7	1.9	490	1 Y04127	Mammalian cytochrome P	2.79e+01
12	7	1.9	490	1 R89862	Cytochrome P450 2C9 cl	2.79e+01
13	7	1.9	490	1 R93168	Human cytochrome P450	2.79e+01
14	7	1.9	490	1 R81465	Human cytochrome P450	2.79e+01
15	7	1.9	490	1 W64070	Human cytochrome P450	2.79e+01
16	7	1.9	490	1 W64070	Human cytochrome P450	2.79e+01
17	7	1.9	490	1 W64070	Human cytochrome P450	2.79e+01
18	7	1.9	490	1 R89865	Human cytochrome P450	2.79e+01
19	7	1.9	490	1 R73374	Human auxiliary cytoch	2.79e+01
20	7	1.9	490	1 R93181	Human cytochrome P450	2.79e+01
21	7	1.9	490	1 R73361	Human cytochrome P450	2.79e+01
22	7	1.9	494	1 W03763	Interleukin binding fa	2.79e+01
23	7	1.9	540	1 W03762	Interleukin binding fa	2.79e+01

24	7	1.9	589	1 R94386	Human neural cell prot	2.79e+01
25	7	1.9	589	1 R94389	Mouse neural cell prot	2.79e+01
26	7	1.9	771	1 W60590	Rat hepatocyte nuclear	2.79e+01
27	7	1.9	898	1 W14777	Granulocyt virus infec	2.79e+01
28	7	1.9	1285	1 W72972	Drosophila melanogaste	2.79e+01
29	7	1.9	1299	1 R86304	Drosophila patched pro	2.79e+01
30	7	1.9	1374	1 W69753	Herpes simplex virus-1	2.79e+01
31	7	1.9	1374	1 W72225	HSV-2 strain SB5 Conti	2.79e+01
32	7	1.9	1384	1 W72224	HSV-2 strain SB5 Conti	2.79e+01
33	7	1.9	1396	1 W72039	HSV-2 strain SB5 Conti	2.79e+01
34	7	1.9	1396	1 W72117	HSV-2 strain SB5 Conti	2.79e+01
35	6	1.6	63	1 Y12007	Human 5' EST secreted	3.36e+02
36	6	1.6	73	1 Y01362	Modified K11 RNA polym	3.36e+02
37	6	1.6	83	1 Y01331	Modified K11 RNA polym	3.36e+02
38	6	1.6	83	1 W97225	Rat type ligand polytpe	3.36e+02
39	6	1.6	108	1 Y03361	HBV specific single st	3.36e+02
40	6	1.6	297	1 Y03363	HBV specific single st	3.36e+02
41	6	1.6	482	1 Y05915	Amino acid sequence of	3.36e+02
42	6	1.6	506	1 W75451	Human wild type tub pr	3.36e+02
43	6	1.6	945	1 W88428	Chlamydia pneumoniae s	3.36e+02
44	6	1.6	1528	1 W99895	Mouse multidrug resist	3.36e+02
45	6	1.6	1829	1 Y07242	Actin-filament binding	3.36e+02

ALIGNMENTS

RESULT 1  
 ID R94349 standard; Protein; 499 AA.  
 AC R94349; 1996 (first entry)  
 DT 20-JUN-1996  
 DE Rat cytochrome P450 2C11.  
 KW Cytochrome P450 2C11; arachidonic acid epoxigenase;  
 KW salt-induced hypertension; diagnosis; transgenic animal;  
 KW therapy.  
 OS Rattus sp.  
 PN WO9610074-A1.  
 PD 04-APR-1996.  
 PF 27-SEP-1995; UI3051.  
 PR 28-SEP-1994; US-314601.  
 PA (UVVA-) UNIV VANDERBILT.  
 PI Capdevilla J, Karara A, Makita K;  
 DR WPI: 96-200913/20.  
 PT Diagnosis and treatment of salt-induced hypertension - using prods.  
 PT developed using mutant P450 2C11 arachidonic acid epoxigenase gene.  
 PS Disclosure; Page 40-42; 55pp; English.  
 CC Normally expressed rat cytochrome P450 arachidonic acid  
 CC epoxigenase (PAAE) has the amino acid sequence given in R94349.  
 CC Mutations in PAAE and its human homologue are associated with a  
 CC predisposition to salt-induced hypertension (SIH). Products and  
 CC treatment of SIH, in screening assays for effective drugs, and in the  
 CC breeding of transgenic animals.  
 CC N.B. the translated sequence of the PAAE gene (T13412) given in  
 CC the specification does not correspond to this rat PAAE amino acid  
 CC sequence.  
 SQ Sequence 499 AA;

Query Match 2.2%; Score 8; DB 1; Length 499;  
 Best Local Similarity 100.0%; Pred. No. 1.90e+00;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 13 SLLLSLW 20

QY 245 SLLLSLW 252

RESULT 2  
 ID W93223 standard; Protein; 40 AA.  
 AC W93223;  
 DT 27-MAY-1999 (first entry)  
 DE Human cytochrome P450 2C10 peptide fragment #1.  
 KW Cytochrome P450; detergent; enzyme inhibitor; non-ionic; ionic;  
 KW purification; drug oxidation; steroid; carcinogen; pesticide; human.

OS Homo sapiens.  
 PN US586157-A.  
 PD 23-MAR-1999.  
 PF 10-FEB-1994; 194981.  
 PR 10-FEB-1994; US-194981.  
 PA (UYVA-) UNIV VANDERBILT.  
 PI Gillam EMJ, Guengerich FP, Guo Z, Sandhu P;  
 DR WPI; 99-228609/19.  
 PT Purifying recombinant cytochrome P450 - by utilising novel  
 combinations of detergents and enzyme inhibitors  
 PS Example 2; Fig 10; 91pp; English.  
 CC This invention describes a recombinant cytochrome P450 protein which  
 is purified from a host cell culture using a combination of detergents  
 and enzyme inhibitors. The method comprises (a) fractionating the host  
 cells to prepare their membranes, (b) adding a non-ionic detergent to  
 the membranes in a concentration of 0.8% to 2% (w/v) in a  
 detergent:protein ratio of between 4:1 to 10:1, (c) adding an ionic  
 detergent to the membranes in a concentration of 0.4% to 0.8% (w/v) in a  
 detergent:protein ratio of between 2:1 to 4:1 (d) centrifuging the  
 membrane-detergent mixture to remove insoluble materials and (e)  
 purifying the protein through a diethylaminoethyl-beaded column, then  
 through a carboxymethyl-beaded column, and finally through a  
 hydroxylapatite column. The method is used to purify cytochrome P450  
 proteins which are responsible for catalysing the oxidation of drugs,  
 steroids, carcinogens and pesticides. The method simplifies the  
 purification of P450 proteins, by the use of improved expression vectors  
 and novel detergent combinations.  
 SQ Sequence 40 AA;

Query Match 1.9%; Score 7; DB 1; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 LLLLSLW 20  
 |||||  
 Qy 246 LLLLSLW 252

RESULT 3  
 ID W93224 standard; Protein; 40 AA.  
 AC W93224;  
 DE 27-MAY-1999 (first entry)  
 DT Human cytochrome P450 2C10 peptide fragment #2.  
 KW Cytochrome P450; detergent; enzyme inhibitor; non-ionic; ionic;  
 purification; drug oxidation; steroid; carcinogen; pesticide; human.  
 OS Homo sapiens.  
 PN US586157-A.  
 PD 23-MAR-1999.  
 PF 10-FEB-1994; 194981.  
 PR 10-FEB-1994; US-194981.  
 PA (UYVA-) UNIV VANDERBILT.  
 PI Gillam EMJ, Guengerich FP, Guo Z, Sandhu P;  
 DR WPI; 99-228609/19.  
 PT Purifying recombinant cytochrome P450 - by utilising novel  
 combinations of detergents and enzyme inhibitors  
 PS Example 2; Fig 10; 91pp; English.  
 CC This invention describes a recombinant cytochrome P450 protein which  
 is purified from a host cell culture using a combination of detergents  
 and enzyme inhibitors. The method comprises (a) fractionating the host  
 cells to prepare their membranes, (b) adding a non-ionic detergent to  
 the membranes in a concentration of 0.8% to 2% (w/v) in a  
 detergent:protein ratio of between 4:1 to 10:1, (c) adding an ionic  
 detergent to the membranes in a concentration of 0.4% to 0.8% (w/v) in a  
 detergent:protein ratio of between 2:1 to 4:1 (d) centrifuging the  
 membrane-detergent mixture to remove insoluble materials and (e)  
 purifying the protein through a diethylaminoethyl-beaded column, then  
 through a carboxymethyl-beaded column, and finally through a  
 hydroxylapatite column. The method is used to purify cytochrome P450  
 proteins which are responsible for catalysing the oxidation of drugs,  
 steroids, carcinogens and pesticides. The method simplifies the  
 purification of P450 proteins, by the use of improved expression vectors  
 and novel detergent combinations.  
 SQ Sequence 40 AA;

Query Match 1.9%; Score 7; DB 1; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 LLLLSLW 20  
 |||||  
 Qy 246 LLLLSLW 252

RESULT 4  
 ID W93226 standard; Protein; 41 AA.  
 AC W93226;  
 DE 27-MAY-1999 (first entry)  
 DT Human cytochrome P450 2C10 peptide fragment #4.  
 KW Cytochrome P450; detergent; enzyme inhibitor; non-ionic; ionic;  
 purification; drug oxidation; steroid; carcinogen; pesticide; human.  
 OS Homo sapiens.  
 PN US586157-A.  
 PD 23-MAR-1999.  
 PF 10-FEB-1994; 194981.  
 PR 10-FEB-1994; US-194981.  
 PA (UYVA-) UNIV VANDERBILT.  
 PI Gillam EMJ, Guengerich FP, Guo Z, Sandhu P;  
 DR WPI; 99-228609/19.  
 PT Purifying recombinant cytochrome P450 - by utilising novel  
 combinations of detergents and enzyme inhibitors  
 PS Example 2; Fig 10; 91pp; English.  
 CC This invention describes a recombinant cytochrome P450 protein which  
 is purified from a host cell culture using a combination of detergents  
 and enzyme inhibitors. The method comprises (a) fractionating the host  
 cells to prepare their membranes, (b) adding a non-ionic detergent to  
 the membranes in a concentration of 0.8% to 2% (w/v) in a  
 detergent:protein ratio of between 4:1 to 10:1, (c) adding an ionic  
 detergent to the membranes in a concentration of 0.4% to 0.8% (w/v) in a  
 detergent:protein ratio of between 2:1 to 4:1 (d) centrifuging the  
 membrane-detergent mixture to remove insoluble materials and (e)  
 purifying the protein through a diethylaminoethyl-beaded column, then  
 through a carboxymethyl-beaded column, and finally through a  
 hydroxylapatite column. The method is used to purify cytochrome P450  
 proteins which are responsible for catalysing the oxidation of drugs,  
 steroids, carcinogens and pesticides. The method simplifies the  
 purification of P450 proteins, by the use of improved expression vectors  
 and novel detergent combinations.  
 SQ Sequence 41 AA;

Query Match 1.9%; Score 7; DB 1; Length 41;  
 Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 15 LLLLSLW 21  
 |||||  
 Qy 246 LLLLSLW 252

RESULT 5  
 ID Y12020 standard; Protein; 72 AA.  
 AC Y12020;  
 DE 18-JUN-1999 (first entry)  
 DT Human 5' EST secreted protein SEQ ID NO: 333.  
 KW Human; secreted protein; EST; expressed sequence tag; diagnosis;  
 forensic; gene therapy; chromosome mapping; signal peptide;  
 upstream regulatory sequence; cytokine activity; cell proliferation;  
 differentiation; haematopoiesis regulation; tissue growth regulation;  
 reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;  
 thrombolytic; anti-inflammatory; tumour inhibition.  
 OS Homo sapiens.  
 PN W0906554-A.  
 PD 11-FEB-1999.  
 PR 31-JUL-1998; IB1238.  
 PR 01-AUG-1997; US-905134.  
 PI (GEST) GENSET.  
 Duciart A, Dumas Milne Edwards J, Lacroix B;

DR WPI: 99-153784/13.  
 DR N-PSDB: X40853.  
 PT New nucleic acids encoding human secreted proteins - obtained from  
 PT cDNA libraries prepared from kidney, fetal kidney, dystrophic  
 PT muscle, muscle and heart tissue  
 PS Claim 34; Page 472; 622pp; English.  
 CC X40826 to X41093 represent 5' expressed sequence tags (ESTs) for human  
 CC secreted proteins, and encode the proteins given in Y01602 and  
 CC Y11994 to Y12260, respectively. The proteins given represent the signal  
 CC peptide and an N-terminal fragment of a secreted protein. The nucleic  
 CC acid sequences can be used for producing secreted human gene products.  
 CC They can also be used to develop products for diagnosis and therapy.  
 CC The proteins obtained may have cytokine activity, cell  
 CC proliferation/differentiation activity, haematopoiesis regulating  
 CC activity, tissue growth regulating activity, reproductive hormone  
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and  
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, tumour inhibition activity or other activities. The products  
 CC can be used in forensic, gene therapy and chromosome mapping procedures.  
 CC The sequences can also be used for obtaining corresponding promoter  
 CC sequences. The nucleic acids encoding the signal peptide can be used  
 CC for directing extracellular secretion of a polypeptide or the insertion  
 CC of a polypeptide into a membrane, or importing a polypeptide into  
 CC a cell.  
 SQ Sequence 72 AA;

Query Match 1.9%; Score 7; DB 1; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 12 GGSLLQLP 18  
 QY 338 GGSLLQLP 344  
 |||||

RESULT 6  
 ID P50297 standard; Protein; 105 AA.  
 AC P50297;  
 DT 03-DEC-1991 (first entry)  
 DE Human epidermal growth factor (EGF).  
 KW Epidermal growth factor; ss.  
 OS Homo sapiens.  
 PN W08500369-A.  
 PD 31-JAN-1985.  
 PF 02-JUL-1984; U01050.  
 PR 05-JUL-1983; US-511372.  
 PA (CHIR-) CHIRON CORP.  
 PI Graeme BI;  
 DR WPI: 85-038094/06.  
 DR N-PSDB: N50343.  
 PT New human DNA sequence encoding epidermal growth factor - useful  
 PT for prodn. of EGF and related polypeptide(s).  
 PS Disclosure; Page 12-12a; 21pp; English.  
 CC Human EGF is a potent mitogen for a variety of cells and a potent  
 CC inhibitor of gastric acid secretion.  
 SQ Sequence 105 AA;

Query Match 1.9%; Score 7; DB 1; Length 105;  
 Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 94 LLLLSLW 100  
 QY 246 LLLLSLW 252  
 |||||

RESULT 7  
 ID W60650 standard; Protein; 112 AA.  
 AC W60650;  
 DT 12-OCT-1998 (first entry)  
 DE Human C-C chemokine DGWCC.  
 KW DGWCC; DNAX groin wound expressed CC chemokine; cytokine; human;  
 KW immune system; cancer; cell proliferation; therapy; diagnosis.

OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Peptide 1..24  
 FT /label= Sig\_peptide  
 FT Protein 25..112  
 FT /label= Mat\_protein  
 FT /note= "Claim 2"  
 FT W09823750-A2.  
 PN 04-JUN-1998.  
 PD 26-NOV-1997; U21092.  
 PR 05-DEC-1996; US-761071.  
 PR 27-NOV-1996; US-031805.  
 PA (SCHE ) SCHERING CORP.  
 PI Hedrick JA, Morales J, Vicari A, Zlotnik A;  
 DR WPI: 98-322730/28.  
 DR N-PSDB: V38294.  
 DT Vic-1 and DGWCC chemokines - useful for developing products for  
 PT treating abnormal physiology or development, e.g. cancerous or  
 PT degenerative conditions  
 PS Claim 2; Page 62; 71pp; English.  
 CC This polypeptide comprises novel human DNAX groin wound expressed  
 CC CC chemokine (DGWCC), the mature portion of which is claimed. The  
 CC DGWCC amino acid sequence was deduced from a cDNA clone (see V38294).  
 CC Also claimed is novel human DNAX Vic-1 (DVia-1) (see W60649) mature  
 CC polypeptide, as well as expression vectors and host cells. DGWCC  
 CC and DVic-1 play a role in the regulation or development of neuronal  
 CC or haematopoietic cells, e.g. lymphoid cells, which affect  
 CC immunological responses. They can be used in the treatment of  
 CC conditions associated with abnormal physiology or development,  
 CC including abnormal proliferation, e.g. cancerous conditions or  
 CC degenerative conditions. Abnormal proliferation, regeneration,  
 CC degeneration, and atrophy may be modulated by appropriate  
 CC therapeutic treatment using products of the invention. The products  
 CC can also be used for detection, diagnosis and drug screening.  
 SQ Sequence 112 AA;

Query Match 1.9%; Score 7; DB 1; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 SLLLSL 15  
 QY 245 SLLLSL 251  
 |||||

RESULT 8  
 ID W44082 standard; Protein; 131 AA.  
 AC W44082;  
 DT 12-MAY-1998 (first entry)  
 DE Human secreted protein AK296.  
 KW Human; secreted protein; ATCC 98026; cytokine; immunomodulation;  
 KW cell proliferation; differentiation; regulation.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Misc\_difference 16  
 FT /label= Unspecified  
 FT /note= "encoded by MTA"  
 FT Misc\_difference 17  
 FT /label= Unspecified  
 FT /note= "encoded by TAG a stop codon"  
 FT Misc\_difference 37  
 FT /label= Unspecified  
 FT /note= "encoded by TAG a stop codon"  
 FT Misc\_difference 40  
 FT /label= Unspecified  
 FT /note= "encoded by TGA a stop codon"  
 FT Misc\_difference 107  
 FT /label= Unspecified  
 FT /note= "encoded by TGA a stop codon"  
 FT Misc\_difference 111  
 FT /label= Unspecified  
 FT /note= "encoded by NGT"  
 FT Misc\_difference 112

FT /label= Unspecified  
FT /note= "encoded by TAG a stop codon"  
PN  
PD WO9739123-A2.  
PD 23-OCT-1997.  
PF 14-APR-1997; U06139.  
PF 18-APR-1996; US-634325.  
PA (GENY ) GENETICS INST INC.  
PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA,  
PI Spaulding V;  
PI WPI; 97-526460/48.  
DR N-PSDB: V02147.  
DR New secreted proteins encoded clones present in ATCC 98026 -  
PT possibly having cytokine, cell proliferation/differentiation  
PT regulating, immunomodulating and many other activities  
PS Claim 24; Page 87-88; 139pp; English.  
CC The present sequence represents a novel human secreted protein deposited  
CC under accession number ATCC 98026. The secreted protein can be used to  
CC determine biological activity, to raise antibodies, as tissue markers,  
CC to isolate cognate ligands or receptors, to identify agents that  
CC modulate their interactions and as nutritional supplements. It may also  
CC have a very wide range of biological activities although no evidence  
CC for any is provided in the specification. Typical of these are cytokine,  
CC cell proliferation/differentiation modulating activity or induction of  
CC other cytokines; immunostimulating/immunosuppressant activities (e.g.  
CC for treating human immunodeficiency virus infection, cancer, autoimmune  
CC diseases and allergy); regulation of haematopoiesis (e.g. for treating  
CC anaemia or as adjunct to chemotherapy); stimulation of growth of bone,  
CC cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds,  
CC periodontal disease, neurological diseases stroke, fibrosis); inhibition  
CC or stimulation of follicle stimulating hormone (for control of  
CC fertility); chemotactic and chemokinetic activities (e.g. for treating  
CC infections, tumours); haemostatic or thrombolytic activity (e.g. for  
CC treating haemophilia, cardiac infarction etc.); anti-inflammatory  
CC activity (e.g. for treating septic shock, Crohn's disease); as  
CC antimicrobials; for treating psoriasis or other hyperproliferative  
CC disease; for regulation of metabolism, behaviour, and many others. Also  
CC contemplated is the use of the corresponding nucleic acid in gene  
CC therapy procedures.  
SQ Sequence 131 AA;

Query Match 1.9%; Score 7; DB 1; Length 131;  
Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 25 SLLLSL 31  
QY 245 SLLLSL 251

RESULT 9  
ID W27646 standard; Protein; 131 AA.  
AC W27646;  
DT 14-MAY-1998 (first entry)  
DE Secreted protein AK296.  
KW Human; secreted protein; research; treatment; AK296.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Misc\_difference 16  
FT /label= unknown  
FT /note= "encoded by MTA"  
FT Misc\_difference 18  
FT /label= unknown  
FT /note= "encoded by TAG"  
FT Misc\_difference 37  
FT /label= unknown  
FT /note= "encoded by TAG"  
FT Misc\_difference 40  
FT /label= unknown  
FT /note= "encoded by TGA"  
FT Misc\_difference 107  
FT /note= "encoded by TGA"  
FT Misc\_difference 111  
FT /label= unknown

FT Misc\_difference 112 /note= "encoded by NGR"  
FT /label= unknown  
FT /note= "encoded by TAG"  
PN  
PD WO9739122-A2.  
PD 23-OCT-1997.  
PF 11-APR-1997; U06042.  
PF 12-APR-1996; US-631184.  
PA (MURO-) MURO PHARM INC.  
PI Theoharides TC;  
PI WPI; 97-526459/48.  
DR N-PSDB: T88064.  
DR Human and murine secreted proteins - useful to research or treat  
PT diseases or disorders related to their function  
PT Claim 23; Pages 87-88; 140pp; English.  
PS The present sequence is a human secreted protein, which may  
CC have nutritional uses, or cytokine and cell  
CC proliferation/differentiation, immune stimulating or suppressing,  
CC haematopoiesis regulating, tissue growth, activin/inhibin,  
CC chemotactic/chemokinetic, haemostatic and thrombolytic,  
CC receptor/ligand, anti-inflammatory or tumour inhibition activities.  
CC It can also be used to research or treat diseases/disorders related  
CC to its function.  
CC The partial cDNA clone AP162 was 1st isolated from a human adult  
CC placenta cDNA library. The partial cDNA clones AM931, AM610, AM340,  
CC AM282, AK647, AK533 and AK296 were 1st isolated from a human  
CC foetal kidney cDNA library. The partial cDNA clones H617 and B89  
CC were 1st isolated from a human peripheral blood monocyte cell (Th1  
CC or Th2) cDNA library. The partial cDNA clone AW191 was 1st isolated  
CC from a human ovary (PA-1 teratocarcinoma) cDNA library. The partial  
CC cDNA clones AT211, AT205 and AT319 were 1st isolated from a human  
CC lymphocyte and dendritic cell cDNA library. The partial cDNA clones  
CC AS34 and AS32 were 1st isolated from a human foetal brain cDNA  
CC library. The partial cDNA clone AR260 was 1st isolated from a human  
CC adult retina cDNA library. The partial cDNA clones K640 and K39  
CC were 1st isolated from a murine bone marrow (stromal cell line  
CC FCM-4) cDNA library.  
SQ Sequence 131 AA;

Query Match 1.9%; Score 7; DB 1; Length 131;  
Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 25 SLLLSL 31  
QY 245 SLLLSL 251

RESULT 10  
ID W55091 standard; Protein; 393 AA.  
AC W55091;  
DT 02-OCT-1998 (first entry)  
DE Streptococcus pneumoniae Sp0038 protein.  
KW Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;  
KW detection; pneumonia; otitis media; meningitis.  
OS Streptococcus pneumoniae.  
FH Key Location/Qualifiers  
FT Misc\_difference 273  
FT /label= unknown  
FT /note= "encoded by GNN"  
FT Misc\_difference 301  
FT /label= unknown  
FT /note= "encoded by NAC"  
PN  
PD WO9818930-A2.  
PD 07-MAY-1998.  
PF 30-OCT-1997; U19422.  
PR 31-OCT-1996; US-029960.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Choi GH, Hromockyj A, Johnson LS, Kunsch CA;  
DR WPI; 98-272224/24.  
DR N-PSDB: V27352.  
PT Nucleic acid encoding antigenic peptide(s) from Streptococcus  
PT pneumoniae - or their epitope-containing fragments, useful in

PT protective or therapeutic vaccines, and for diagnosis  
 PS Claim 11; Page 60; 118pp; English.  
 CC The present sequence represents a protein from Streptococcus pneumoniae.  
 CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein  
 CC can be useful in vaccines for inducing protective antibodies against  
 CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.  
 CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid  
 CC are used to detect Streptococcus infection (by usual hybridisation or  
 CC amplification methods), also for isolating Streptococcus genes or their  
 CC allelic variants. The protein can be used similarly for detecting specific  
 CC antibodies in standard immunoassays especially for diagnosing or  
 CC monitoring infections. Antibodies which bind the protein are used to  
 CC detect corresponding antigens, to purify the protein and for passive  
 CC immunisation (optionally coupled to a toxin). Vaccines are administered,  
 CC e.g. by injection, orally or through the skin, typically at 0.01-1000  
 CC (especially 10-300)  $\mu$ g/ml per dose.  
 SQ Sequence 393 AA;

Query Match 1.9%; Score 7; DB 1; Length 393;  
 Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 72 LGQGGA 78  
 QY 21 LGQGGA 27

RESULT 11  
 ID Y04127 standard; Protein; 490 AA.

AC Y04127;  
 DT 11-JUN-1999 (first entry)  
 DE Mammalian cytochrome P450 protein CYP2C9.  
 KW Bacterial; mammalian; cytochrome P450; chimeric; fusion protein;  
 KW oxidase; hydrocarbon; carbon-hydrogen bond; hydroxylating;  
 KW bioremediation; environmental pollutant.  
 OS Mammalia  
 FN WO9908612-A1.  
 PD 25-FEB-1999.  
 PF 17-AUG-1998; UI6979.  
 PR 20-AUG-1997; US-056754.  
 PA (UYRP ) UNIV ROCHESTER.  
 PI Jones JP, Shimoji M;  
 WI; 99-190131/16.  
 DR N-PSDB; X19925.  
 PT New P450 fusion proteins - comprising a portion of a bacterial  
 PT cytochrome P450 protein and a portion of a mammalian cytochrome P450  
 PT protein  
 PS Disclosure; Page 9-10; 51pp; English.  
 CC The present invention describes a fusion proteins comprising a portion  
 CC of a bacterial cytochrome P450 protein and also a portion of a mammalian  
 CC cytochrome P450 protein. The fusion protein can oxidise hydrocarbons or  
 CC any compound having a carbon-hydrogen bond. The fusion protein can be  
 CC used for hydroxylating a compound to be oxidised. It can also be used in  
 CC the bioremediation of an environmental pollutant. Since the fusion  
 CC protein is soluble, it can be subject to structural elucidation by X-ray  
 CC crystallography for designing functional proteins. It can be readily  
 CC expressed in soil bacteria to facilitate bioremediation. The present  
 CC sequence represents mammalian cytochrome P450 protein CYP2C9 from the  
 CC present invention.  
 SQ Sequence 490 AA;

Query Match 1.9%; Score 7; DB 1; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 LLLLSLW 20  
 QY 246 LLLLSLW 252

RESULT 12  
 ID R89862 standard; Protein; 490 AA.

AC R89862;

DT 09-SEP-1996 (first entry)  
 DE Cytochrome P450 2C9 clone 25.  
 KW Cytochrome P450 2C19; human; liver; PCR; primer; detection; CYP2C19;  
 KW stereospecific S-mephenytoin 4'-hydroxylase activity; polymorphism.  
 OS Homo sapiens.  
 FH key Location/Qualifiers  
 FT misc\_difference 359  
 FT /note= "Variable position"  
 FT WO9530766-A1.  
 PD 16-NOV-1995.  
 PF 08-MAY-1995; U05744.  
 PR 06-MAY-1994; US-238821.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PI De Morais SMF, Goldstein JA, Romkes-Sparks M;  
 WI; 96-077257/08.  
 DR N-PSDB; T11378.  
 DT New isolated cytochrome P450 2C subfamily member - used for  
 PT identifying drugs metabolised by S-mephenytoin 4'-hydroxylase  
 PT activity and to develop other screening assays  
 PS Example 2; Page 93-94; 16pp; English.  
 CC The sequences given in R89862-65 represent allelic variants of cytochrome  
 CC P450 2C9. The majority of clones isolated from liver S33 coded for 2C9.  
 CC Of the 50 clones encoding 2C9, only two allelic variants were found. 39  
 CC of the 2C9 clones were identical with clone 65, and 11 were identical  
 CC with clone 25. Clones 25 and 65 are identical in the 5' and  
 CC 3'-noncoding regions but contained 2 single base changes at positions  
 CC 1075 and 1425. One of these base changes was conservative but the other  
 CC results in one amino acid difference at position 359, Ile to Leu. The  
 CC clones 29c and 6b differ by one nucleotide in the coding region, Thr to Met  
 CC at position 1154, which results in a single amino acid change, Thr to Met  
 CC at position 385. Clone 29c has a very long, 198 bp, 5'-noncoding region  
 CC and a polyadenylation signal 21 bases from the poly-A tail. Clone 6b  
 CC has an unusually long 3'-noncoding region containing three possible  
 CC polyadenylation signals with no poly-A tail. The differences in the 3'  
 CC non-coding regions could represent alternate splicing, allelic  
 CC variants, or possibly separate genes. These clones are designated as  
 CC allelic variants of 2C18 because they differ by only one base in the  
 CC coding region. They are most similar to 2C9 (82% amino acid homology)  
 CC and 2C19 (81% amino acid homology).  
 SQ Sequence 490 AA;

Query Match 1.9%; Score 7; DB 1; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 LLLLSLW 20  
 QY 246 LLLLSLW 252

RESULT 13  
 ID R93168 standard; Protein; 490 AA.

AC R93168;  
 DT 11-OCT-1996 (first entry)  
 DE Human cytochrome P450 molecular species 2C9 protein.  
 KW Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;  
 KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;  
 KW evaluation; safety; fusion protein; metabolite; detoxification;  
 KW carcinogenic.  
 OS Homo sapiens.  
 FN J08056695-A.  
 PD 05-MAR-1996.  
 PF 15-JUL-1994; 164184.  
 PR 20-JUL-1993; JP-201120.  
 PR 30-JUL-1993; JP-208279.  
 PR 17-JUN-1994; JP-136053.  
 PA (SUMO ) SUMITOMO CHEM CO LTD.  
 DR WI; 96-182311/19.  
 DT N-PSDB; T28381.  
 PT Novel method for the evaluation of the safety of a cpd. - using a  
 PT human cytochrome P450 and yeast NADPH reductase to determine whether  
 PT the analyte cpd. is detoxified or metabolised to a carcinogen  
 PS Example 1; Page 20-22; 74pp; Japanese.

CC This is the amino acid sequence of the human cytochrome P450 molecular  
 CC species 2C9 protein. The corresp. gene was amplified from a human liver  
 CC derived cDNA library as 2 fragments of 0.9 and 0.6 kb using primers  
 CC T26925-8. The prod. was cloned into the yeast expression vector pRAH5N  
 CC to generate plasmid p2C9 for prodn. of the cytochrome only or into the  
 CC vector pAHRP to generate the plasmid p2C9R for co-prodn. with the yeast  
 CC NADPH-P450 reductase. The sequence is placed under control of the yeast  
 CC ADH gene promoter and terminator.  
 CC The vectors are used in a method for evaluating the safety of a cpd. by  
 CC reacting the test cpd. with recombinantly produced human cytochrome P450  
 CC mol. species 1A2 (T28380), 2C9, 2E1 (T28382), 3A4 (T28383) or their  
 CC variants (T28384-98) together with yeast NADPH-P450 reductase (either as  
 CC a fused protein or as a cell extract) and analysing the resultant  
 CC metabolite. The cpd. is considered "safe" if it is detoxified or not  
 CC rendered carcinogenic or "unsafe" if it is not detoxified or is  
 CC metabolised to a carcinogenic cpd.  
 SQ Sequence 490 AA;

Query Match 1.9%; Score 7; DB 1; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 14 LLLLSLW 20  
 |||||  
 Qy 246 LLLLSLW 252

RESULT 14  
 ID R81465 standard; Protein; 490 AA.  
 AC R81465;  
 DT 01-AUG-1996 (first entry)  
 DE Human derived cytochrome P4502C9.  
 KW Human derived cytochrome; P4502C9; commercial cDNA library; yeast;  
 KW transfection; recombinant production; expression vector; mammal;  
 KW immunisation; sensitisation; antibody; determination; detection;  
 KW non-cross reactive.  
 OS Homo sapiens.  
 PN J08027196-A.  
 PD 30-JAN-1996.  
 PF 13-JUL-1994; 161551.  
 PR 13-JUL-1994; JP-161551.  
 PA (SUMO ) SUMITOMO CHEM CO LTD.  
 DR WPI; 96-136337/14.  
 DR N-PSDB; T17404.  
 PT Antibody recognising human derived cytochrome P4502C9 - allows  
 PT specific detection of cytochrome P450 species in humans  
 PS Example 1; Pages 11-13; 13pp; Japanese.  
 CC The present sequence is the human derived cytochrome (HDC)  
 CC P4502C9, which was obt'd. from a commercial cDNA library. Yeast  
 CC were transfected with an expression vector contg. the HDC cDNA,  
 CC cultured and then disrupted to give a microsomal fraction. The  
 CC HDC was purified from the fraction, and used to immunise and  
 CC sensitise a mammal. Blood was drawn from the mammal, and an  
 CC anti-HDC antibody isolated. The antibody obt'd. recognises HDC  
 CC P4502C9, partic. at a serum dilution rate of 1:10000, and is  
 CC substantially without cross reaction to other HDC P450 spp. .  
 SQ Sequence 490 AA;

Query Match 1.9%; Score 7; DB 1; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 14 LLLLSLW 20  
 |||||  
 Qy 246 LLLLSLW 252

RESULT 15  
 ID W64073 standard; Protein; 490 AA.  
 AC W64073;  
 DT 06-OCT-1998 (first entry)  
 DE Human cytochrome P450 2C9 clone 65 protein.  
 KW Cytochrome P450; drug screening; S-mephenytoin 4'-hydroxylase;

KW identification; mutagenic; carcinogenic; cytotoxic; haemoprotein;  
 KW xenobiotic; environmental pollutant.  
 OS Homo sapiens.  
 PN US5786191-A.  
 PD 28-JUL-1998.  
 PF 22-FEB-1994; 201118.  
 PR 22-FEB-1994; US-201118.  
 PR 09-APR-1992; US-864962.  
 PA (GOLD/) GOLDSTEIN J A.  
 PA (ROMK/) ROMKES-SPARKS M.  
 PI Goldstein JA, Romkes-sparks M;  
 DR WPI; 98-436528/37.  
 PT Screening for drugs metabolised by cytochrome P450 - for identifying  
 PT mutagenic, carcinogenic, or cytotoxic compounds  
 PS Example 2; Column 51-54; 63pp; English.  
 CC This sequence represents a human cytochrome P450 2C9 polypeptide isolated  
 CC from clone 65. This polypeptide is used in a method to screen for a drug  
 CC that is metabolised by a cytochrome P450 having S-mephenytoin  
 CC 4'-hydroxylase activity. The protein can also be used to identify a  
 CC mutagenic, carcinogenic or cytotoxic compound. Cytochrome P450 are a  
 CC large family of haemoprotein enzymes capable of metabolising xenobiotics  
 CC such as drugs carcinogens and environmental pollutants as well as  
 CC endobiotics such as steroids, fatty acids and prostaglandins.  
 SQ Sequence 490 AA;

Query Match 1.9%; Score 7; DB 1; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 2.79e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 LLLLSLW 20  
 |||||  
 Qy 246 LLLLSLW 252

Search completed: Wed May 10 14:39:54 2000  
 Job time : 16 secs.

US-09-376-430-2.ra1

Thu May 11 06:50:38 2000

\*\*\*\*\*  
 M P S R E L H  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 14:40:11 2000; MasPar time 50.94 Seconds  
 94.365 Million cell updates/sec  
 Tabular output not generated.

Title: >US-09-376-430-2  
 Description: (1-371) from US09376430A.pep  
 Perfect Score: 371  
 Sequence: 1 MGRVLVLMGAAVFLGGWMA.....DVTIGGFTVMNDRSYVAL 371

Scoring table: TABLE unitprotable  
 Gap 60

Searched: 131253 seqs, 12956647 residues  
 Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: a-issued  
 1:5A\_COMB 2:5B\_COMB 3:PCT\_COMB 4:backfiles1

Statistics: Mean 2.793; Variance 0.718; scale 3.889  
 Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description	Pred. No.
1	8	2.2	US-08-314-2	Sequence 2, Applicatio	1.16e+00
2	8	2.2	PCT-US95-1	Sequence 2, Applicatio	1.16e+00
3	7	1.9	US-08-194-1	Sequence 15, Applicati	1.62e+01
4	7	1.9	US-08-194-2	Sequence 14, Applicati	1.62e+01
5	7	1.9	US-08-194-3	Sequence 17, Applicati	1.62e+01
6	7	1.9	US-07-885-1	Sequence 13, Applicati	1.62e+01
7	7	1.9	PCT-US92-0	Sequence 15, Applicati	1.62e+01
8	7	1.9	US-08-330-1	Sequence 13, Applicati	1.62e+01
9	7	1.9	US-08-440-1	Sequence 13, Applicati	1.62e+01
10	7	1.9	US-07-847-1	Sequence 15, Applicati	1.62e+01
11	7	1.9	US-08-456-1	Sequence 15, Applicati	1.62e+01
12	7	1.9	US-08-456-2	Sequence 15, Applicati	1.62e+01
13	7	1.9	US-08-419-1	Sequence 13, Applicati	1.62e+01
14	7	1.9	PCT-US95-0	Sequence 3, Applicati	1.62e+01
15	7	1.9	US-08-238-1	Sequence 9, Applicati	1.62e+01
16	7	1.9	US-08-238-2	Sequence 13, Applicati	1.62e+01
17	7	1.9	US-08-201-1	Sequence 3, Applicati	1.62e+01
18	7	1.9	PCT-US95-0	Sequence 13, Applicati	1.62e+01
19	7	1.9	US-08-201-2	Sequence 3, Applicati	1.62e+01
20	7	1.9	US-08-238-3	Sequence 3, Applicati	1.62e+01
21	7	1.9	PCT-US95-0	Sequence 9, Applicati	1.62e+01
22	7	1.9	US-07-906-1	Sequence 33, Applicati	1.62e+01
23	7	1.9	US-07-906-2	Sequence 33, Applicati	1.62e+01

24	7	1.9	US-07-906-1	Sequence 2, Applicatio	1.62e+01
25	7	1.9	PCT-US95-1	Sequence 4, Applicatio	1.62e+01
26	7	1.9	PCT-US95-1	Sequence 2, Applicatio	1.62e+01
27	7	1.9	US-08-317-1	Sequence 2, Applicatio	1.62e+01
28	7	1.9	US-08-317-2	Sequence 4, Applicatio	1.62e+01
29	7	1.9	US-08-742-1	Sequence 2, Applicatio	1.62e+01
30	7	1.9	PCT-US95-1	Sequence 6, Applicatio	1.62e+01
31	7	1.9	US-08-540-1	Sequence 6, Applicatio	1.62e+01
32	6	1.6	US-08-540-2	Patent No. 5175383	1.88e+02
33	6	1.6	US-07-949-1	Sequence 9, Applicatio	1.88e+02
34	6	1.6	US-07-949-2	Patent No. 5175383	1.88e+02
35	6	1.6	US-08-462-1	Sequence 2, Applicatio	1.88e+02
36	6	1.6	US-08-736-1	Sequence 6, Applicatio	1.88e+02
37	6	1.6	US-08-736-2	Sequence 24, Applicati	1.88e+02
38	6	1.6	US-08-407-1	Sequence 2, Applicatio	1.88e+02
39	6	1.6	US-08-188-1	Patent No. 5231789	1.88e+02
40	6	1.6	US-08-714-1	Sequence 27, Applicati	1.88e+02
41	6	1.6	US-08-413-1	Patent No. 5422248	1.88e+02
42	6	1.6	US-08-413-2	Sequence 9, Applicatio	1.88e+02
43	6	1.6	US-08-332-1	Sequence 42, Applicati	1.88e+02
44	6	1.6	US-08-241-1	Sequence 2, Applicatio	1.88e+02
45	6	1.6	US-08-276-1	Sequence 2, Applicatio	1.88e+02

ALIGNMENTS

RESULT 1 STANDARD: PRT; 500 AA.

ID US-08-314-601-2  
 XX AC xxxxxx

Sequence 2, Application US/08314601

Sequence 2, Application US/08314601

Patent No. 5834293

GENERAL INFORMATION:

APPLICANT: Capdevilla, Jorge H.

TITLE OF INVENTION: CYTOCHROME P450 ARACHIDONIC ACID

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: NEEDLE & ROSENBERG, P.C.

STREET: 127 Peachtree Street, Suite 1200

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/314,601

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Perryman, David G.

REGISTRATION NUMBER: 33,438

REFERENCE/DOCKET NUMBER: 2200.024

TELECOMMUNICATION INFORMATION:

TELEPHONE: (404) 688-0770

TELEFAX: (404) 688-9880

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 500 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SQ SEQUENCE 500 AA; 57181 MW; 1320568 CN;

Query Match 2.2%; Score 8; DB 2; Length 500;  
Best Local Similarity 100.0%; Pred. No. 1.16e+00;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 13 SLLLSLW 20  
|||||

QY 245 SLLLSLW 252

RESULT 2

ID PCT-US95-13051-2 STANDARD; PRT; 500 AA.  
XX AC xxxxxx

Sequence 2, Application PC/TUS9513051

Sequence 2, Application PC/TUS9513051  
GENERAL INFORMATION:

APPLICANT: VANDERBILT UNIVERSITY  
TITLE OF INVENTION: CYTOCHROME P450 ARACHIDONIC ACID  
TITLE OF INVENTION: EPOXYGENASE GENETIC MUTATION ASSOCIATED WITH HYPERTENSION  
NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
STREET: 127 Peachtree Street, Suite 1200  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13051  
FILING DATE: 27-SEP-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/314,601  
FILING DATE: 9/28/94

ATTORNEY/AGENT INFORMATION:  
NAME: Perryman, David G.

REGISTRATION NUMBER: 33,438  
REFERENCE/DOCKET NUMBER: 2200.024

TELEPHONE: (404) 688-0770  
TELEFAX: (404) 688-9880

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 500 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SQ SEQUENCE 500 AA; 57181 MW; 1320568 CN;

Query Match 2.2%; Score 8; DB 3; Length 500;  
Best Local Similarity 100.0%; Pred. No. 1.16e+00;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 13 SLLLSLW 20  
|||||

QY 245 SLLLSLW 252

RESULT 3

ID US-08-194-981E-15 STANDARD; PRT; 40 AA.  
XX

AC xxxxxx

Sequence 15, Application US/08194981E  
Patent No. 5886157  
GENERAL INFORMATION:

APPLICANT: GUENGERICH, F. Peter  
APPLICANT: GUO, Zuyu

APPLICANT: SANDHU, Punam

APPLICANT: GILLAM, Elizabeth M. J.

TITLE OF INVENTION: EXPRESSION AND PURIFICATION OF

TITLE OF INVENTION: HUMAN

TITLE OF INVENTION: CYTOCHROME P450

NUMBER OF SEQUENCES: 68

CORRESPONDENCE ADDRESS:

ADDRESSEE: NEEDLE & ROSENBERG, P.C.

STREET: Suite 1200, 127 Peachtree Street, NE

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303-1811

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/194,981E

FILING DATE: February 10, 1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Elizabeth Selby

REGISTRATION NUMBER: 38,298

REFERENCE/DOCKET NUMBER: 22000.0022

TELEPHONE: (404) 688-0770

TELEFAX: (404) 688-9880

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 40 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

FRAGMENT TYPE: N-terminal

SQ SEQUENCE 40 AA; 4232 MW; 9710 CN;

Query Match 1.9%; Score 7; DB 2; Length 40;  
Best Local Similarity 100.0%; Pred. No. 1.62e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 LLLLSLW 20  
|||||

QY 246 LLLLSLW 252

RESULT 4

ID US-08-194-981E-14 STANDARD; PRT; 40 AA.  
XX

AC xxxxxx

Sequence 14, Application US/08194981E

Sequence 14, Application US/08194981E

Patent No. 5886157

GENERAL INFORMATION:

APPLICANT: GUENGERICH, F. Peter





CC FILING DATE: 18-MAY-1992  
CC CLASSIFICATION: 530  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Coruzzi, Laura A.  
CC REGISTRATION NUMBER: 30,742  
CC REFERENCE/DOCKET NUMBER: 5624-174  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 212-790-9090  
CC TELEFAX: 212-869-9741  
CC TELEX: 66141 PENNIE  
CC INFORMATION FOR SEQ ID NO: 13:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 88 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC SEQUENCE 88 AA; 10157 MW; 43797 CN;  
SQ

Query Match 1.9%; Score 7; DB 2; Length 88;  
Best Local Similarity 100.0%; Pred. No. 1.62e-01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 71 LLLSLW 77  
| | | | |  
Qy 246 LLLSLW 252

RESULT 7  
ID PCT-US92-04295A-15 STANDARD; PRT; 91 AA.  
XX AC  
XX xxxxxx  
XX  
XX  
XX  
DE Sequence 15, Application PC/TUS9204295A  
CC Sequence 15, Application PC/TUS9204295A  
CC GENERAL INFORMATION:  
CC APPLICANT: Genentech, Inc.  
CC TITLE OF INVENTION: Structure, Production and Use of  
CC TITLE OF INVENTION: Heregulin  
CC NUMBER OF SEQUENCES: 30  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Genentech, Inc.  
CC STREET: 460 Point San Bruno Blvd  
CC CITY: South San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94080  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: patin (Genentech)  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US92/04295A  
CC FILING DATE: 19920521  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 07/705256  
CC FILING DATE: 24-MAY-1991  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 07/765212  
CC FILING DATE: 25-SEP-1991  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 07/790801  
CC FILING DATE: 08-NOV-1991  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 07/847743  
CC FILING DATE: 06-MAR-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Hensley, Max D.

CC REGISTRATION NUMBER: 27,043  
CC REFERENCE/DOCKET NUMBER: 712P4  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 415/225-1994  
CC TELEFAX: 415/952-9881  
CC TELEX: 910/371-7168  
CC INFORMATION FOR SEQ ID NO: 15:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 91 amino acids  
CC TYPE: AMINO ACID  
CC TOPOLOGY: linear  
CC SEQUENCE 91 AA; 10463 MW; 46593 CN;  
SQ

Query Match 1.9%; Score 7; DB 3; Length 91;  
Best Local Similarity 100.0%; Pred. No. 1.62e-01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 76 LLLSLW 82  
| | | | |  
Qy 246 LLLSLW 252

RESULT 8  
ID US-08-330-161-13 STANDARD; PRT; 91 AA.  
XX AC  
XX xxxxxx  
XX  
XX  
XX  
DE Sequence 13, Application US/08330161  
XX Sequence 13, Application US/08330161  
CC Patent No. 5834229  
CC GENERAL INFORMATION:  
CC APPLICANT: Vandlen, Richard  
CC APPLICANT: Holmes, William  
CC TITLE OF INVENTION: Structure, Production and Use of Heregulin 2 Ligands  
CC NUMBER OF SEQUENCES: 17  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Genentech, Inc.  
CC STREET: 460 Point San Bruno Blvd  
CC CITY: South San Francisco  
CC STATE: California  
CC COUNTRY: USA  
CC ZIP: 94080  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: patin (Genentech)  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/330,161  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 08/035430  
CC FILING DATE: 22-MAR-1993  
CC APPLICATION NUMBER: 07/705256  
CC FILING DATE: 24-MAY-1991  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Adler, Carolyn R.  
CC REGISTRATION NUMBER: 32,324  
CC REFERENCE/DOCKET NUMBER: 712C1  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 415/225-2614  
CC TELEFAX: 415/952-9881  
CC TELEX: 910/371-7168  
CC INFORMATION FOR SEQ ID NO: 13:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 91 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC SEQUENCE 91 AA; 10463 MW; 46593 CN;  
SQ

Thu May 11 06:50:38 2000

Query Match 1.9%; Score 7; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 1.62e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 76 LLLSLW 82  
| | | | | | | |  
QY 246 LLLSLW 252

RESULT 9  
ID US-08-440-401-13 STANDARD; PRT; 91 AA.  
XX  
AC xxxxxx

Sequence 13, Application US/08440401  
Patent No. 5856110  
GENERAL INFORMATION:  
APPLICANT: Vandlen, Richard L.  
TITLE OF INVENTION: Structure, Production and Use of  
TITLE OF INVENTION: Heregulin 2 Ligands  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patin (Genentech)  
CURRENT APPLICATION DATA:  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/330161  
FILING DATE: 23-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/035430  
FILING DATE: 22-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/705256  
FILING DATE: 24-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 712C3  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 91 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE 91 AA; 10463 MW; 46593 CN;

Query Match 1.9%; Score 7; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 1.62e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 76 LLLSLW 82  
| | | | | | | |

QY 246 LLLSLW 252

RESULT 10  
ID US-07-847-743B-15 STANDARD; PRT; 91 AA.  
XX  
AC xxxxxx

Sequence 15, Application US/07847743B  
Patent No. 5367060  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
TITLE OF INVENTION: Structure, Production and Use of  
TITLE OF INVENTION: Heregulin  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/847,743B  
FILING DATE: 19920306  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/705256  
FILING DATE: 24-MAY-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/765212  
FILING DATE: 25-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/790801  
FILING DATE: 08-NO. 5367060-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hensley, Max D.  
REGISTRATION NUMBER: 27,043  
REFERENCE/DOCKET NUMBER: 712P3  
TELEPHONE: 415/266-1489  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 91 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
SEQUENCE 91 AA; 10463 MW; 46593 CN;

Query Match 1.9%; Score 7; DB 1; Length 91;  
Best Local Similarity 100.0%; Pred. No. 1.62e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 76 LLLSLW 82  
| | | | | | | |  
QY 246 LLLSLW 252

RESULT 11  
ID US-08-456-201-15 STANDARD; PRT; 91 AA.  
XX  
AC xxxxxx  
XX



Thu May 11 06:50:38 2000

Sequence 13, Application US/08419878B  
Patent No. 5853206  
GENERAL INFORMATION:  
APPLICANT: Mandlen, Richard L.  
TITLE OF INVENTION: Antibodies Specific For Heregulin 2-alpha  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genattech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/419,878B  
FILING DATE: 11-APR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/330161  
FILING DATE: 25-OCT-1994  
APPLICATION NUMBER: 08/035430  
FILING DATE: 22-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/705256  
FILING DATE: 24-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0712C2D1  
TELEPHONE: 650/225-1994  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 91 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
SEQUENCE 91 AA; 10463 MW; 46593 CN;  
Query Match 1.9%; Score 7; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 1.62e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 76 LLLLSLW 82  
QY 246 LLLLSLW 252  
RESULT 14  
ID PCT-US95-05744-3 STANDARD; PRT; 490 AA.  
XX  
AC  
XX  
XX  
DT  
XX  
DE  
Sequence 3, Application PC/TUS9505744  
GENERAL INFORMATION:  
APPLICANT: GOLDSTEIN, Joyce A.  
APPLICANT: ROMKES-SPARKS, Marjorie  
APPLICANT: DE MORAIS, Sonia M.F.  
TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN  
CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT  
TITLE OF INVENTION: OF S-MEPHNYTOIN METABOLISM  
NUMBER OF SEQUENCES: 61

Sequence 9, Application US/08238821B  
Patent No. 5812120  
GENERAL INFORMATION:  
APPLICANT: GOLDSTEIN, Joyce A.  
APPLICANT: ROMKES-SPARKS, Marjorie  
APPLICANT: DE MORAIS, Sonia M.F.  
TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN  
CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT  
TITLE OF INVENTION: MEPHNYTOIN METABOLISM  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Crew LLP  
STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/05744  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA: US 08/238,821  
APPLICATION NUMBER:  
FILING DATE: 06-MAY-1994  
PRIOR APPLICATION DATA: US 08/201,118  
APPLICATION NUMBER:  
FILING DATE: 22-FEB-1994  
PRIOR APPLICATION DATA: US 07/864,962  
APPLICATION NUMBER:  
FILING DATE: 09-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 15280-192-1-1  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 490 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
SEQUENCE 490 AA; 55627 MW; 1288532 CN;  
Query Match 1.9%; Score 7; DB 3; Length 490;  
Best Local Similarity 100.0%; Pred. No. 1.62e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 14 LLLLSLW 20  
QY 246 LLLLSLW 252  
RESULT 15  
ID US-08-238-821B-9 STANDARD; PRT; 490 AA.  
XX  
AC  
XX  
XX  
DT  
DE  
Sequence 9, Application US/08238821B  
Patent No. 5812120  
GENERAL INFORMATION:  
APPLICANT: GOLDSTEIN, Joyce A.  
APPLICANT: ROMKES-SPARKS, Marjorie  
APPLICANT: DE MORAIS, Sonia M.F.  
TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN  
CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT  
TITLE OF INVENTION: MEPHNYTOIN METABOLISM  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco

CC STATE: California  
 CC COUNTRY: US  
 CC ZIP: 94111  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/238,821B  
 CC FILING DATE: 06-MAY-1994  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/201,118  
 CC FILING DATE: 22-FEB-1994  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07/864,962  
 CC FILING DATE: 09-APR-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Liebeschuetz, Joe  
 CC REGISTRATION NUMBER: 37,505  
 CC REFERENCE/DOCKET NUMBER: 15280-192110US  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (650) 326-2400  
 CC TELEFAX: (650) 326-2422  
 CC INFORMATION FOR SEQ ID NO: 9:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 490 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 490 AA; 55627 MW; 1288173 CN;

Query Match 1.9%; Score 7; DB 2; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 1.62e-01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 LLLLSLW 20  
 QY 245 LLLLSLW 252

Search completed: Wed May 10 14:41:11 2000  
 Job time : 60 secs.



```

##cross-references EMBL:AF049028; NID:g3098570; PID:g3098571
##experimental_source cv. Topas, microspore
SUMMARY #length 282 #molecular-weight 32089 #checksum 8665

Query Match 2.2%; Score 8; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 6.37e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 15 LLLLLSLW 22
Qy 245 LLLLLSLW 252

RESULT 3
ENTRY #type complete
TITLE cytochrome P450 2C4 - rabbit
CONTAINS oxidoreductase (EC 1.-.-.-)
ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic
rabbit
DATE 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change
ACCESSIONS 11-Jun-1999
REFERENCE A26731; B34257
#authors Johnson, E.F.; Barnes, H.J.; Griffin, K.J.; Okino, S.; Tukey, R.H.
#journal J. Biol. Chem. (1987) 262:5918-5923
#title Characterization of a second gene product related to rabbit cytochrome P-450 1.
#cross-references MUID:87194795
#accession A26731
##molecule_type DNA
##residues 1-487 #label JOH
##cross-references GB:J02716; NID:g164924; PID:AAA31216.1; PID:g164925
REFERENCE A94658
#authors Zhao, J.; Chan, G.; Govind, S.; Bell, P.; Kemper, B.
#journal DNA Cell Biol. (1990) 9:37-48
#title Structure of 5' regions and expression of phenobarbital-inducible rabbit cytochrome P450IIC genes.
#cross-references MUID:90197893
#accession B34257
##molecule_type DNA
##residues 1-10 #label ZH2
##cross-references GB:M74200
CLASSIFICATION #superfamily human cytochrome P450 CYP2D6; cytochrome P450
homology
KEYWORDS chromoprotein; electron transfer; heme; iron; monooxygenase;
oxidoreductase; transmembrane protein
FEATURE 432
#binding_site heme iron (Cys) (axial ligand) #status
predicted
SUMMARY #length 487 #molecular-weight 55386 #checksum 1184

Query Match 2.2%; Score 8; DB 2; Length 487;
Best Local Similarity 100.0%; Pred. No. 6.37e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 LLLLLSLW 21
Qy 246 LLLLLSLW 253

RESULT 4
ENTRY #type complete
TITLE cytochrome P450 2C1 - rabbit
CONTAINS oxidoreductase (EC 1.-.-.-)
ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic
rabbit
DATE 20-Sep-1984 #sequence_revision 16-Feb-1996 #text_change
ACCESSIONS 11-Jun-1999
REFERENCE A00181; A34257
#authors Leighton, J.K.; DeBrunner-Vossbrinck, B.A.; Kemper, B.
#journal Biochemistry (1984) 23:204-210

```

```

#title Isolation and sequence analysis of three cloned cDNAs for
rabbit liver proteins that are related to rabbit cytochrome
P-450 (form 2), the major phenobarbital-inducible form.
#cross-references MUID:84128536
#accession A00181
##molecule_type mRNA
##residues 11-490 #label LE2
##cross-references EMBL:K01522; NID:g164914; PID:AAA31211.1;
PID:g164915
##note the authors translated the codon CAA for residue 48 as
Lys, GAT for residue 133 as Asn, CTG for residue 291
as Val, AGT for residue 292 as Thr, GTG for residue
439 as Ala, and GCC for residue 441 as Val

REFERENCE A94658
#authors Zhao, J.; Chan, G.; Govind, S.; Bell, P.; Kemper, B.
#journal DNA Cell Biol. (1990) 9:37-48
#title Structure of 5' regions and expression of
phenobarbital-inducible rabbit cytochrome P450IIC genes.
#cross-references MUID:90197893
#accession A34257
##molecule_type DNA
##residues 1-24 #label ZHA
##cross-references GB:M74199
GENETICS CYP2C1
#superfamily human cytochrome P450 CYP2D6; cytochrome P450
homology
CLASSIFICATION chromoprotein; electron transfer; endoplasmic reticulum;
heme; iron; monooxygenase; oxidoreductase; transmembrane
protein
KEYWORDS
FEATURE 435
#binding_site heme iron (Cys) (axial ligand) #status
predicted
SUMMARY #length 490 #molecular-weight 55614 #checksum 4387

Query Match 2.2%; Score 8; DB 1; Length 490;
Best Local Similarity 100.0%; Pred. No. 6.37e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 LLLLLSLW 21
Qy 246 LLLLLSLW 253

RESULT 5
ENTRY #type complete
TITLE laurate omega-minus-1 hydroxylase (EC 1.14.14.-) cytochrome
P450 2C2
CONTAINS cytochrome P450 PBC2
ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic
rabbit
DATE 20-Sep-1984 #sequence_revision 16-Feb-1996 #text_change
ACCESSIONS 11-Jun-1999
REFERENCE A27718; A00182; S15587
#authors Inai, Y.; Komori, M.; Sato, R.
#journal Biochemistry (1988) 27:80-88
#title Comparison of primary structures deduced from cDNA nucleotide
sequences for various forms of liver microsomal cytochrome
P-450 from phenobarbital-treated rabbits.
#cross-references MUID:88163622
#accession A27718
##molecule_type mRNA
##residues 1-490 #label IMA
##experimental_source clone Hp2
REFERENCE A90484
#authors Leighton, J.K.; DeBrunner-Vossbrinck, B.A.; Kemper, B.
#journal Biochemistry (1984) 23:204-210
#title Isolation and sequence analysis of three cloned cDNAs for
rabbit liver proteins that are related to rabbit cytochrome
P-450 (form 2), the major phenobarbital-inducible form.
#cross-references MUID:84128536
#accession A00182

```



US-09-376-430-2.rpr

Thu May 11 06:50:39 2000

```

##molecule_type mRNA
##residues 12-470, 'L', 472-490 #label LEI
##cross-references EMBL:R01521; NID:g164912; PIDN:AAA31210.1;
PID:g164913

REFERENCE
#authors Govind, S.; Bell, P.A.; Kemper, B.
#journal DNA (1986) 5:371-382
#title Structure of genes in the cytochrome P-450Pbc subfamily:
phenobarbital-inducible family.
#cross-references MUID:87053173
#accession S15587
#status preliminary
##molecule_type DNA
##residues 1-22 #label GOV
##cross-references EMBL:M14955; NID:g164908; PIDN:AAA31208.1;
PID:g164909

GENETICS
#gene CYP2C2
#superfamily human cytochrome P450 CYP2D6; cytochrome P450
homology
#classification
#keywords chromoprotein; electron transfer; endoplasmic reticulum;
heme; iron; monooxygenase; oxidoreductase; transmembrane
protein
#feature 435
#binding_site heme iron (Cys) (axial ligand) #status
predicted
#length 490 #molecular-weight 55791 #checksum 1914
#query Match 2.2%; Score 8; DB 1; Length 490;
#best Local Similarity 100.0%; Pred. No. 6.37e-02;
#matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 LLLLSLWK 21
QY 246 LLLLSLWK 253

RESULT 6
ENTRY A26921 #type complete
TITLE cytochrome P450, phenobarbital-inducible, hepatic - rabbit
CONTAINS oxidoreductase (EC 1.-.-.-)
ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic
rabbit
DATE 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change
11-Jun-1999
ACCESSIONS A26921
REFERENCE Imai, Y.
#authors J. Biochem. (1987) 101:1129-1139
#journal Cytochrome P-450 related to P-450-4 from
#title phenobarbital-treated rabbit liver: molecular cloning of
cDNA and characterization of cytochrome P-450 obtained by
its expression in yeast cells.
#cross-references MUID:88007455
#accession A26921
##molecule_type mRNA
##residues 1-490 #label IMA
##cross-references GB:D00190; NID:g217741; PIDN:BA00130.1;
PID:d1000575; PID:g217742

GENETICS
#gene CYP2C14
#superfamily human cytochrome P450 CYP2D6; cytochrome P450
homology
#classification
#keywords chromoprotein; electron transfer; heme; iron; liver;
monooxygenase; oxidoreductase; transmembrane protein
#feature 435
#binding_site heme iron (Cys) (axial ligand) #status
predicted
#length 490 #molecular-weight 55720 #checksum 2584
#query Match 2.2%; Score 8; DB 2; Length 490;
#best Local Similarity 100.0%; Pred. No. 6.37e-02;

##molecule_type mRNA
##residues 1-500 #label BIA
##cross-references EMBL:U33173; NID:g944945; PIDN:AAB02144.1;
PID:g944946

REFERENCE
#authors Biagini, C.; Cellier, C.
#journal Arch. Biochem. Biophys. (1996) 326:298-305
#title CDNA-directed expression of two allelic variants of
cytochrome P450 2C11 using COS1 and Sf21 insect cells.
#cross-references MUID:96190638
#accession S62779
##molecule_type mRNA
##residues 1-500 #label BIA
##cross-references EMBL:U33173
#classification #superfamily human cytochrome P450 CYP2D6; cytochrome P450
homology
#keywords chromoprotein; heme; iron; transmembrane protein
#feature 435
#binding_site heme iron (Cys) (axial ligand) #status
predicted
#length 500 #molecular-weight 57091 #checksum 9715
#query Match 2.2%; Score 8; DB 2; Length 500;
#best Local Similarity 100.0%; Pred. No. 6.37e-02;
#matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 13 LLLLSLWK 20
QY 245 LLLLSLWK 252

RESULT 8
ENTRY A26685 #type complete
TITLE cytochrome P450 2C11 - rat
ALTERNATE_NAMES cytochrome P450 M-1; P450-16alpha; P450-UT-A; testosterone
16alpha-dehydrogenase (EC 1.1.1.-) CYP2C11
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change
11-Jun-1999
ACCESSIONS A26685; A29421; A60782; A60783; S26818; S44310; S62779
REFERENCE A26685
#authors Yoshioka, H.; Morohashi, K.; Sogawa, K.; Miyata, T.;
Kawajiri, K.; Hirose, T.; Inayama, S.; Fujii-Kuriyama, Y.;
Omura, T.
#journal J. Biol. Chem. (1987) 262:1706-1711
#title Structural analysis and specific expression of microsomal
cytochrome P450(M-1) mRNA in male rat livers.
#cross-references MUID:87103321
#accession A26685
##molecule_type mRNA
##residues 1-500 #label YOS
##cross-references GB:J02657; NID:g203867; PIDN:AAA41062.1; PID:g203868
A29421
REFERENCE

```

```

#authors Morishima, N.; Yoshioka, H.; Higashi, Y.; Sogawa, K.;
#journal Fujii-Kuriyama, Y.
#title Biochemistry (1987) 26:8279-8285
#cross-references Gene structure of cytochrome P-450 (M-1) specifically
#accession expressed in male rat liver.
#molecule_type MUID:88163490
#residues A29421
#cross-references 1-328,'H',330-500 ##label MOR
#note the single nucleotide difference in this sequence was
#note verified by bidirectional sequencing and may represent
#note allelic polymorphism
REFERENCE
#authors A60782
#journal Stroem, A.; Mode, A.; Zaphiropoulos, P.; Nilsson, A.G.;
#title Morgan, E.; Gustafsson, J.A.
#cross-references Acta Endocrinol. (1988) 118:314-320
#accession Cloning and pretranslational hormonal regulation of
#molecule_type testosterone 16alpha-hydroxylase (P-450-16alpha) in male
#residues rat liver.
#status not compared with conceptual translation
#molecule_type MUID:88266578
#residues A60782
#cross-references not compared with conceptual translation
#accession MUID:88266578
#status not compared with conceptual translation
#molecule_type mRNA
#residues 1-500 ##label STR
REFERENCE
#authors A60783
#journal Zaphiropoulos, P.G.; Mode, A.; Stroem, A.; Husman, B.;
#title Anderson, G.; Gustafsson, J.A.
#cross-references Acta Med. Scand. Suppl. (1988) 723:161-167
#accession Sequence and regulation of two growth-hormone-controlled,
#molecule_type sex-specific isozymes of cytochrome P-450 in rat liver,
#residues P-450-15beta and P-450-16alpha.
#cross-references MUID:88279070
#accession A60783
#status nucleic acid sequence not shown; not compared with
#molecule_type conceptual translation
#residues 1-500 ##label ZAP
REFERENCE
#authors Matsumoto, T.; Emi, Y.; Kawabata, S.; Omura, T.
#journal J. Biochem. (1986) 100:1359-1371
#title Purification and characterization of three male-specific and
#cross-references one female-specific forms of cytochrome P-450 from rat
#accession liver microsomes.
#molecule_type MUID:87137396
#residues S26818
#cross-references not compared with conceptual translation
#accession MUID:87137396
#molecule_type protein
#residues 1-30 ##label MAT
REFERENCE
#authors S44310
#journal Stroehm, A.; Equichi, H.; Mode, A.; Tollet, P.; Stroemstedt,
#title P.; Gustafsson, J.
#cross-references submitted to the EMBL Data Library, May 1994
#accession Characterization of the proximal promoter and two silencer
#molecule_type elements in the CYP2C gene expressed in rat liver.
#residues S44310
#status preliminary
#molecule_type DNA
#residues 1-11,'L',13-56 ##label ST2
#cross-references EMBL:X79081; NID:9483487; PIDN:CAA55686.1;
#accession PID:9483488
REFERENCE
#authors S62779
#journal Biagini, C.; Cellier, C.
#title Arch. Biochem. Biophys. (1996) 326:298-305
#cross-references cDNA-directed expression of two allelic variants of
#accession cytochrome P450 2C11 using COS1 and SF21 insect cells.
#molecule_type MUID:96190658
#residues S62779
#status preliminary; not compared with conceptual translation
#molecule_type mRNA
#residues 1-8,112-120;183-191 ##label B1A
#experimental_source strain Wistar
#comment This cytochrome P450 is expressed specifically in adult male rat
#comment liver.

```

US-09-376-430-2.rpr

Thu May 11 06:50:39 2000

```

#cross-references MUID:93041757
#accession S26663 preliminary
#status preliminary
#molecule_type DNA
#residues 1-66 #label AN2
#cross-references EMBL:X61371; NID:g36716; PID:g36717
REFERENCE
S41125
Andreadis, A.; Nisson, P.E.; Kosik, K.S.; Watkins, P.C.
Nucleic Acids Res. (1993) 21:2217-2221
The exon trapping assay partly discriminates against
alternatively spliced exons.
#journal
#title
#cross-references MUID:93275752
#accession S41125
#status preliminary
#molecule_type DNA
#residues 1-66 #label AN2
GENETICS
#gene GDB:MAPT; MTBT1
#cross-references GDB:119434; OMIM:157140
#map_position 17q21-17q21
CLASSIFICATION
#superfamily microtubule-associated protein tau; MAP2/tau
#repeat homology
#length 66 #checksum 4316
SUMMARY
Query Match 1.9%; Score 7; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.84e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 42 PSSPKHV 48
| | | | |
QY 119 PSSPKHV 125

RESULT 11
ENTRY A56755 #type fragment
TITLE cytochrome P450 2C24, major splice form - rat (fragment)
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change
11-Jun-1999
A56755
A56755
Zaphiropoulos, P.G.
Biochem. Biophys. Res. Commun. (1993) 192:778-786
Differential expression of cytochrome P450 2C24 transcripts
in rat kidney and prostate: evidence indicative of
alternative and possibly trans splicing events.
#cross-references MUID:93249456
#accession A56755
#status preliminary
#molecule_type DNA; mRNA
#residues 1-70 #label ZAP
#cross-references GB:S59652; NID:g300021; PIDN:AAB26469.1; PID:g300023
#experimental_source Sprague-Dawley
#note sequence extracted from NCBI backbone (NCBIN:130999,
NCBIN:131007, NCBI:P131009)
CLASSIFICATION
#superfamily human cytochrome P450 CYP2D6; cytochrome P450
homology
#alternative splicing; heme; transmembrane protein
#length 70 #checksum 4154
KEYWORDS
SUMMARY
Query Match 1.9%; Score 7; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 4.84e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 14 LLLSLW 20
| | | | |
QY 246 LLLSLW 252

RESULT 12
ENTRY C72234 #type complete
TITLE hypothetical protein - Thermotoga maritima (strain MSB8)
ORGANISM #formal_name Thermotoga maritima

DATE 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change
11-Jun-1999
C72234
A72200
Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson,
R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Nelson,
W.C.; Ketchum, K.A.; McDonald, L.; Utterback, T.R.; Malek,
J.A.; Liner, K.D.; Garrett, M.M.; Stewart, A.M.; Cotton,
M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
Heidelberg, J.; Sutton, G.G.; Fleischmann, R.D.; White, O.;
Salzberg, S.L.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Nature (1999) 399:323-329
Evidence for lateral gene transfer between Archaea and
Bacteria from genome sequence of Thermotoga maritima.
#journal
#title
#cross-references MUID:99287316
#accession C72234
#status preliminary
#molecule_type DNA
#residues 1-120 #label ARN
#cross-references GB:AE001804; GB:AE000512; NID:g4982160; PID:g4982170;
TIGR:TM1599
#experimental_source strain MSB8
GENETICS
#gene TM1599
#length 120 #molecular-weight 14025 #checksum 8931
SUMMARY
Query Match 1.9%; Score 7; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 4.84e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 75 VSLLLS 81
| | | | |
QY 244 VSLLLS 250

RESULT 13
ENTRY S63305 #type complete
TITLE probable membrane protein YNL324w - yeast (Saccharomyces
cerevisiae)
ALTERNATE_NAMES hypothetical protein N0332
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 27-Apr-1996 #sequence_revision 03-May-1996 #text_change
14-Nov-1997
S63305
S63305
Matahi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.
#authors submitted to the Protein Sequence Database, April 1996
#submission
#accession S63305
#molecule_type DNA
#residues 1-131 #label MAF
#cross-references EMBL:271601; NID:g1302439; PID:e239769; PID:g1302441;
MIPS:YNL324w
#experimental_source strain S288C
GENETICS
#map_position 14L
#transmembrane protein
KEYWORDS
FEATURE
58-74
102-118
#domain transmembrane #status predicted #label TM1
#domain transmembrane #status predicted #label TM2
#length 131 #molecular-weight 14855 #checksum 924
SUMMARY
Query Match 1.9%; Score 7; DB 2; Length 131;
Best Local Similarity 100.0%; Pred. No. 4.84e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 58 MVSLLL 64
| | | | |
QY 243 MVSLLL 249

RESULT 14
ENTRY S71220 #type complete
TITLE peptidylprolyl isomerase (EC 5.2.1.8) ROC2 - Arabidopsis

```

thaliana  
 ALTERNATE\_NAMES cytosolic cyclophilin; peptidyl-prolyl cis-trans isomerase;  
 rotamase  
 ORGANISM #formal\_name Arabidopsis thaliana #common\_name mouse-ear  
 cress  
 DATE 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change  
 16-Jul-1999  
 ACCESSIONS S71220  
 REFERENCE S71219  
 #authors Chou, I.T.; Gasser, C.S.  
 #submission submitted to the EMBL Data Library, November 1995  
 #description Characterization of cyclophilin gene family in Arabidopsis  
 thaliana.  
 #accession S71220  
 ##molecule\_type DNA  
 ##residues 1-176 #label CHO  
 ##cross-references EMBL:U04000; NID:g1305456; PIDN:AAB96833.1;  
 PID:g1305457

GENETICS  
 #gene ROC2  
 CLASSIFICATION #superfamily peptidylprolyl isomerase; cyclophilin homology  
 cis-trans-isomerase; cyclosporin A binding  
 KEYWORDS  
 FEATURE 3-171  
 SUMMARY #domain cyclophilin homology #label Cyp  
 #length 176 #molecular-weight 18906 #checksum 4353

Query Match 1.9%; Score 7; DB 2; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 4.84e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 146 VKAMEDV 152  
 |||||  
 Qy 187 VKAMEDV 193

RESULT 15  
 ENTRY  
 TITLE T00192 #type complete  
 ORGANISM hypothetical protein 59 - Staphylococcus aureus phage phi PVL  
 DATE #formal\_name Staphylococcus aureus phage phi PVL  
 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change  
 ACCESSIONS T00192  
 REFERENCE Z14119  
 #authors Kaneko, J.; Kimura, T.; Kawakami, Y.; Tomita, T.; Kamio, Y.  
 #journal Biosci. Biotechnol. Biochem. (1997) 61:1960-1962  
 #title Pantone-valentine leukocidin genes in a phage-like particle  
 isolated from mitomycin C-treated Staphylococcus aureus V8  
 (ATCC 49775).  
 #accession T00192  
 ##status translated from GB/EMBL/DBJ  
 ##molecule\_type DNA  
 ##residues 1-216 #label KAN  
 ##cross-references EMBL:AB009866; NID:d1204727; PID:d1032893  
 SUMMARY #length 216 #molecular-weight 26026 #checksum 5993

Query Match 1.9%; Score 7; DB 2; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 4.84e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 105 RVKAMED 111  
 |||||  
 Qy 186 RVKAMED 192

Search completed: Wed May 10 14:39:20 2000  
 Job time : 27 secs.

\*\*\*\*\*  
M E S R L H  
\*\*\*\*\*  
(TW)

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 14:32:02 2000; MasPar time 99.66 Seconds  
113.371 Million cell updates/sec

Tabular output not generated.

Title: >US-09-376-430-2

Description: (1-371) from US09376430A.pap

Perfect Score: 371

Sequence: 1 MGRLLVLLGAAVFLGGWMA.....DVTIGGTFVMDRSYVAL 371

Scoring table: TABLE unitprotatable

Gap 60

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: swiss-prot38

1:swissprot

Statistics: Mean 3.721; Variance 0.435; scale 8.561

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	8	2.2	480	1	CPCL_RABIT CYTOCHROME P450 2C1 (E	2.36e-02
2	8	2.2	487	1	CPCL_RABIT CYTOCHROME P450 2C4 (E	2.36e-02
3	8	2.2	490	1	CPCL_RABIT CYTOCHROME P450 2C11 (	2.36e-02
4	8	2.2	500	1	CPCL_RAT CYTOCHROME P450 2C11 (	2.36e-02
5	8	2.2	693	1	CAUP_DROME HOMEBOX PROTEIN CAUPO	2.36e-02
6	7	1.9	89	1	FLIQ_AQUAE FLAGELLAR BIOSYNTHETIC	2.24e+00
7	7	1.9	131	1	YN64_YEAST HYPOTHETICAL 14.9 KD P	2.24e+00
8	7	1.9	134	1	IL5_FELCA INTERLEUKIN-5 PRECURSO	2.24e+00
9	7	1.9	221	1	YIEH_ECOLI HYPOTHETICAL 24.7 KD P	2.24e+00
10	7	1.9	310	1	COPD_PSESM COPPER RESISTANCE PROT	2.24e+00
11	7	1.9	339	1	SRG7_CAEEL SRG-7 PROTEIN.	2.24e+00
12	7	1.9	357	1	REPL_ZYGBA TRANS-ACTING FACTOR B	2.24e+00
13	7	1.9	376	1	FOS_FUGRU P55-C-FOS PROTO-ONCOGE	2.24e+00
14	7	1.9	410	1	MDPL_RABIT MICROSOMAL DIPEPTIDASE	2.24e+00
15	7	1.9	436	1	MDHC_PEA MALATE DEHYDROGENASE (	2.24e+00
16	7	1.9	441	1	THCB_RHOER HYPOTHETICAL 48.4 KD P	2.24e+00
17	7	1.9	448	1	YIEH_ECOLI HYPOTHETICAL 54.5 KD P	2.24e+00
18	7	1.9	467	1	YIEH_ECOLI HYPOTHETICAL 54.5 KD P	2.24e+00
19	7	1.9	489	1	CPZ6_CANFA CYTOCHROME P450 2C4 (E	2.24e+00
20	7	1.9	490	1	CPZ6_CANFA CYTOCHROME P450 2C9 (E	2.24e+00
21	7	1.9	490	1	CPCA_HUMAN CYTOCHROME P450 2C10 (	2.24e+00
22	7	1.9	504	1	GNUN_ERWCA ENDOGLUCANASE VI PRECU	2.24e+00
23	7	1.9	505	1	FLJB_SALTU PHASE-2 FLAGELLIN.	2.24e+00

24	7	1.9	556	1	CD19_HUMAN B-LYMPHOCYTE ANTIGEN C	2.24e+00
25	7	1.9	574	1	TYRO_PODAN TYROSINASE (EC 1.14.18	2.24e+00
26	7	1.9	620	1	TYRO_NEUCR INTERLEUKIN ENHANCER-B	2.24e+00
27	7	1.9	655	1	ILF1_HUMAN INTERLEUKIN ENHANCER-B	2.24e+00
28	7	1.9	664	1	2AAA_CAEEL PROBABLE PROTEIN PHOSP	2.24e+00
29	7	1.9	759	1	SC11_YEAST CTR1 SUPPRESSOR PROTEI	2.24e+00
30	7	1.9	902	1	RPB1_DICDI DNA-DIRECTED RNA POLY	2.24e+00
31	7	1.9	1076	1	IF3A_CAEEL PROBABLE EUKARYOTIC TR	2.24e+00
32	7	1.9	1207	1	EGF_HUMAN PRO-EPIDERMAL GROWTH F	2.24e+00
33	7	1.9	1209	1	THR_DROME THREE ROWS PROTEIN.	2.24e+00
34	7	1.9	1286	1	PATC_DROME HEDGEHOG RECEPTOR (PAT	2.24e+00
35	7	1.9	1374	1	VCAP_HSV11 MAJOR CAPSID PROTEIN (	2.24e+00
36	7	1.9	1752	1	RPB1_SCHPO DNA-DIRECTED RNA POLY	2.24e+00
37	6	1.6	219	1	YLP4_ZYMMO HYPOTHETICAL 24.4 KD P	1.19e+02
38	6	1.6	298	1	YNT6_YEAST HYPOTHETICAL 34.3 KD P	1.19e+02
39	6	1.6	370	1	YXER_BAGSU HYPOTHETICAL 38.4 KD P	1.19e+02
40	6	1.6	380	1	YRPA_SHVX HYPOTHETICAL 42.2 KD P	1.19e+02
41	6	1.6	616	1	YGO2_YEAST HYPOTHETICAL 72.6 KD P	1.19e+02
42	6	1.6	626	1	YI17_MICHTU HYPOTHETICAL 68.4 KD P	1.19e+02
43	6	1.6	678	1	YF46_MICHTU HYPOTHETICAL PPE-FAMIL	1.19e+02
44	6	1.6	1111	1	YJEP_HAEIN HYPOTHETICAL PROTEIN H	1.19e+02
45	6	1.6	2476	1	ZAN_PIG ZONADHESIN PRECURSOR.	1.19e+02

ALIGNMENTS

RESULT 1  
ID CPCL\_RABIT STANDARD: PRT: 480 AA.  
AC P00180;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE CYTOCHROME P450 2C1 (EC 1.14.14.1) (CYPI1C1) (P450 PBC1) (FRAGMENT).  
GN CYP2C1.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
NC Euthera; Lagomorpha; Leporidae; Oryctolagus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE; 84128536.  
RX Leighton J.K.; Debrunner-Vossbrink B.A.; Kemper B.;  
RT Isolation and sequence analysis of three cloned cDNAs for rabbit  
RT liver proteins that are related to rabbit cytochrome P-450 (form 2),  
RT the major phenobarbital-inducible form.";  
RL Biochemistry 23:204-210(1984).  
CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN  
CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY IT OXIDIZES A VARIETY  
CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY  
CC ACIDS, AND XENOBIOTICS.  
CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) -> ROH +  
CC OXIDIZED FLAVOPROTEIN + H(2)O.  
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
CC -!- INDUCTION: BY PHENOBARBITAL.  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to: license@isb-sib.ch).  
CC  
CC EMBL; K01522; AAA31211.1;  
CC PIR; A00181; O4RBP.  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
DR PFAM; PF00067; P450; 1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum.  
FT NON\_TER 1  
FT BINDING 425 425 HEME.  
SQ SEQUENCE 480 AA; 54587 MW; A8058C09659612DA CRC64;

Query Match 2.2%; Score 8; DB 1; Length 480;  
 Best Local Similarity 100.0%; Pred. No. 2.36e-02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 LLLLSLWK 11  
 |||||  
 QY 246 LLLLSLWK 253

RESULT 2  
 ID CPCE\_RABIT STANDARD; PRT; 487 AA.  
 AC P11371;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE CYTOCHROME P450 2C4 (EC 1.14.14.1) (CYP1C4) (PROGESTERONE 21-HYDROXYLASE) (P450 PBC4) (P1-88).  
 GN CYP2C4.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 87194795.  
 RA Johnson E.F., Barnes H.J., Griffin K.J., Okino S., Tukey R.H.;  
 RT "Characterization of a second gene product related to rabbit cytochrome P-450 1.1";  
 RL J. Biol. Chem. 262:5918-5923(1987).  
 CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.  
 CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH + OXIDIZED FLAVOPROTEIN + H(2)O.  
 CC -!- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES, AND CARCINOGENS.  
 CC -!- MISCELLANEOUS: THIS PROTEIN DIFFERS FROM OTHER FORMS OF CYTOCHROME P450 IN THAT IT CATALYZES THE 21-HYDROXYLATION OF PROGESTERONE, RESULTING IN THE FORMATION OF DEOXYCORTICOSTERONE.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J02716; AAA31216.1;  
 DR PIR; A26731; A26731.  
 DR PRINTS; PR00385; P450.  
 DR PRINTS; PR00463; EP450I.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR PFAM; PF00067; P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT BINDING 432 432  
 HEME.  
 SQ SEQUENCE 487 AA; 55386 MW; F0E7F5F7512F36A CRC64;

Query Match 2.2%; Score 8; DB 1; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 2.36e-02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 LLLLSLWK 21  
 |||||  
 QY 246 LLLLSLWK 253

RESULT 3  
 ID CPCE\_RABIT STANDARD; PRT; 490 AA.  
 AC P17666;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE CYTOCHROME P450 2C14 (EC 1.14.14.1) (CYP1C14) (PHP3).  
 GN CYP2C14.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88007455.  
 RA Inai Y.;  
 RT "Cytochrome P-450 related to P-4504 from phenobarbital-treated rabbit liver: molecular cloning of cDNA and characterization of cytochrome P-450 obtained by its expression in yeast cells";  
 RL J. Biochem. 101:1129-1139(1987).  
 CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.  
 CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH + OXIDIZED FLAVOPROTEIN + H(2)O.  
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
 CC -!- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES, AND CARCINOGENS.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D00190; BAA0130.1;  
 DR PIR; A26921; A26921.  
 DR PRINTS; PR00385; P450.  
 DR PRINTS; PR00463; EP450I.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR PFAM; PF00067; P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT BINDING 435 435  
 HEME.  
 SQ SEQUENCE 490 AA; 55720 MW; 2572163B0AAA076E CRC64;

Query Match 2.2%; Score 8; DB 1; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 2.36e-02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 LLLLSLWK 21  
 |||||  
 QY 246 LLLLSLWK 253

RESULT 4  
 ID CPCE\_RAT STANDARD; PRT; 500 AA.  
 AC P08683; Q63141; Q64554;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CYTOCHROME P450 2C11 (EC 1.14.14.1) (CYP1C11) (P-450(M-1)) (P450H) (P450-UT-A) (UT-2).  
 GN CYP2C11 OR CYP2C-11.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]

Thu May 11 06:50:39 2000

US-09-376-430-2.rsp

```

RP RX MEDLINE: 87109321.
RA Yoshioka H., Morohashi K., Sogawa K., Miyata T., Kawajiri K.,
RA Hirose T., Inayama S., Fujii-Kuriyama Y., Omura T.;
RT "Structural analysis and specific expression of microsomal cytochrome
RL P-450(M-1) mRNA in male rat livers.";
RL J. Biol. Chem. 262:1706-1711(1987).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE: 88163490.
RA Morishima N., Yoshioka H., Higashi Y., Sogawa K., Fujii-Kuriyama Y.;
RT "Gene structure of cytochrome P-450(M-1) specifically expressed in
RL male rat liver."
RL Biochemistry 26:8279-8285(1987).
RN [3]
RP SEQUENCE FROM N.A.
RA TISSUE=LIVER;
RC MEDLINE: 88266578.
RA Stroem A., Mode A., Zaphiropoulos P., Nilsson A.G., Morgan E.,
RA Gustafsson J.-A.;
RT "Cloning and pretranslational hormonal regulation of testosterone 16
RL alpha-hydroxylase (P-45016 alpha) in male rat liver."
RL Acta Endocrinol. 118:314-320(1988).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR GUNN; TISSUE=LIVER;
RX MEDLINE: 96190658.
RA Biagini C., Cellier C.;
RT "cDNA-directed expression of two allelic variants of cytochrome P450
RL 2C11 using COS1 and SF21 insect cells.";
RL Arch. Biochem. Biophys. 326:298-305(1996).
RN [5]
RP SEQUENCE OF 1-56 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RX MEDLINE: 94347210.
RA Stroem A., Eguchi H., Mode A., Legrauerend C., Tollet P.,
RT Stromstedt P.E., Gustafsson J.-A.;
RT "Characterization of the proximal promoter and two silencer elements
RL in the CYP2C11 gene expressed in rat liver.";
RL DNA Cell Biol. 13:805-819(1994).
CC -!- FUNCTION: METABOLIZE TESTOSTERONE MAINLY IN POSITIONS 2 ALPHA AND
CC 16 ALPHA.
CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC OXIDIZED FLAVOPROTEIN + H(2)O.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND, ENDOPLASMIC RETICULUM.
CC -!- TISSUE SPECIFICITY: LIVER; MALE-SPECIFIC.
CC -!- INDUCTION: CONSTITUTIVELY EXPRESSED.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: J02657; AAA41062.1; --
CC EMBL: M18363; AAA41007.1; --
CC EMBL: M18356; AAA41007.1; JOINED.
CC EMBL: M18357; AAA41007.1; JOINED.
CC EMBL: M18359; AAA41007.1; JOINED.
CC EMBL: M18360; AAA41007.1; JOINED.
CC EMBL: M18361; AAA41007.1; JOINED.
CC EMBL: M18362; AAA41007.1; JOINED.
CC EMBL: U33173; AAB02144.1; --
CC EMBL: X79081; CAAS5686.1; --
CC PIR: A26685; A26685.
CC PIR: A29421; A29421.
CC PIR: A60782; A60782.
CC PIR: A60783; A60783.
CC PRINTS: PR00385; P450.
CC PRINTS: PR00463; EP4501.
CC
CC -----
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
DR PFAM: PF00067; P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Polymorphism.
RT BINDING 435 435 HEME (BY SIMILARITY).
RT VARIANT 4 4 V -> A (IN STRAIN GUNN).
RT VARIANT 116 116 N -> S (IN STRAIN GUNN).
RT VARIANT 187 187 F -> L (IN STRAIN GUNN).
RT CONFLICT 12 12 S -> L (IN REF. 5).
RT CONFLICT 329 329 R -> H (IN REF. 2).
RT CONFLICT 500 AA; 57181 MW; 8DCE0E356D8A5AC3 CRC64;
SQ SEQUENCE 500 AA; 57181 MW; 8DCE0E356D8A5AC3 CRC64;

Query Match 2.2%; Score 8; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 2.36e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 13 SULLLSLW 20
QY 245 SULLLSLW 252

RESULT 5
ID CAUP.DROME STANDARD; PRT; 693 AA.
AC P54269;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HOMEOBOX PROTEIN CAUPOLICAN.
GN CAUP.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Phyllozoa; Drosophilidae; Drosophila.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE: 96180722.
RA Gomez-Skarmeta J.-L., del Corral R.D., de la Calle-Mustienes E.,
RA Ferrer-Marco D., Mordolell J.;
RT "Araucan and caupolican, two members of the novel iroquois complex,";
RT encode homeoproteins that control proneural and vein-forming genes.";
RL Cell 85:95-110(1996)
CC -!- FUNCTION: CONTROLS PRONEURAL AND VEIN FORMING GENES. POSITIVE
CC TRANSCRIPTIONAL CONTROLLER OF AC-SC (ACHAETE-SCUTE). MAY ACT AS AN
CC ACTIVATOR THAT INTERACTS WITH THE TRANSCRIPTIONAL COMPLEX
CC ASSEMBLED ON THE AC AND SC PROMOTERS AND PARTICIPATES IN
CC TRANSCRIPTION INITIATION.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE TALE/IRO FAMILY OF HOMEOBOX PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: X95178; CAA64485.1; W-.
DR HSSP: P02833; 1SAN.
DR FLYBASE: FBgn0015919; caup.
DR PROSITE: PS00027; HOMEOBOX_1; 1.
DR PROSITE: PS50071; HOMEOBOX_2; 1.
DR PFAM: PF00046; Homeobox; 1.
KW Transcription regulation; DNA-binding; Homeobox; Nuclear protein;
KW Developmental protein.
FT DNA_BIND : 226 288 HOMEOBOX (TALE-TYPE).
FT DOMAIN 300 303 POLY-ASP.
FT DOMAIN 405 418 POLY-GLN.
FT DOMAIN 501 516 POLY-GLN.
FT DOMAIN 517 528 POLY-HIS.
FT DOMAIN 565 572 POLY-SER.
FT DOMAIN 513 624 POLY-SER.
FT SEQUENCE 693 AA; 73749 MW; 8E0D6D43C9CDC619 CRC64;
SQ SEQUENCE 693 AA; 73749 MW; 8E0D6D43C9CDC619 CRC64;

```

Query Match 2.2%; Score 8; DB 1; Length 693;  
Best Local Similarity 100.0%; Pred. No. 2.36e-02;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 494 QLPHPLO 501  
QY 342 QLPHPLO 349

RESULT 6  
ID FLIQ\_AQUAE STANDARD; PRT; 89 AA.  
AC O67774;  
DT 15-FEB-2000 (Rel. 39, Created)  
DT 15-FEB-2000 (Rel. 39, Last sequence update)  
DE FLAGELLAR BIOSYNTHETIC PROTEIN FLIQ.  
GN FLIQ OR AQ.1962.  
OS Aquifex aeolicus.  
OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-VFS;  
RX MEDLINE; 98196666.  
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aufay M., Huber R.,  
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
RT "The complete genome of the hyperthermophilic bacterium Aquifex  
aeolicus";  
RL Nature 392:353-358(1998).  
CC -!- FUNCTION: ROLE IN FLAGELLAR BIOSYNTHESIS (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE FLIQ/MOPD/SPAQ FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; AE000765; AAC07731.1;  
DR FRAM; PF01313; Bac\_exp0rt\_3; 1.  
DR PRINTS; PR00952; FLGBIOSNFLQ.  
KW Flagella; Transmembrane.  
FT TRANSMEM 18 38  
FT TRANSWM 49 69  
SQ SEQUENCE 89 AA; 9996 MW; 5F1A2B81AEDC804D CRC64;

Query Match 1.9%; Score 7; DB 1; Length 89;  
Best Local Similarity 100.0%; Pred. No. 2.24e+00;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 63 FLGGWM 69  
QY 13 FLGGWM 19

RESULT 7  
ID YN64\_YEAST STANDARD; PRT; 131 AA.  
AC P53825;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DE HYPOTHETICAL 14.9 KD PROTEIN IN EGT2-KREI INTERGENIC REGION.  
GN YN324W OR N0332.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;  
OC Saccharomycetaceae; Saccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / FY1676;

RX MEDLINE; 95373280.  
RA Maftahi M., Nicaud J.-M., Levesque H., Gaillardin C.;  
RT "Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV  
RT identifies the RPD3, PAS8 and KRE1 loci, five new open reading  
RT frames";  
RL Yeast 11:567-572(1995).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; 246259; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; Z71601; CAA96257.1; -;  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 61 81  
FT TRANSWM 102 122  
SQ SEQUENCE 131 AA; 14855 MW; 2EDCIAD9D3A96ACC CRC64;

Query Match 1.9%; Score 7; DB 1; Length 131;  
Best Local Similarity 100.0%; Pred. No. 2.24e+00;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 58 MVSLLL 64  
QY 243 MVSLLL 249

RESULT 8  
ID IL5\_FELCA STANDARD; PRT; 134 AA.  
AC O77515; O62740;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE INTERLEUKIN-5 PRECURSOR (IL-5) (T-CELL REPLACING FACTOR) (TRF)  
DE (EOSINOPHIL DIFFERENTIATION FACTOR).  
IL5.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 98452719.  
RA Padrid P.A., Qin Y., Wells T.N.C., Solway J., Camoretti-Mercado B.;  
RT "Sequence and structural analysis of feline interleukin-5 CDNA";  
RL Am. J. Vet. Res. 59:1263-1269(1998).  
RN [2]  
RP SEQUENCE OF 12-128 FROM N.A.

RA Harley R., Day M.J., Gruffudd-Jones T.J., Harbour D.A., Helps C.R.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: FACTOR THAT INDUCES TERMINAL DIFFERENTIATION OF LATE-  
CC DEVELOPING B-CELLS TO IMMUNOGLOBULIN SECRETING CELLS  
CC (BY SIMILARITY).

CC -!- SUBUNIT: HOMODIMER, LINKED BY A DISULFIDE BOND (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- SIMILARITY: BELONGS TO THE IL-5 FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; AF025436; AAC64505.1; -;  
DR EMBL; AF051372; AAC05752.1; -;  
DR PRINTS; PR00432; INTERLEUKINS.  
KW Cytokine; Growth factor; Glycoprotein; Signal.  
FT SIGNAL 1 19 BY SIMILARITY.



FT CHAIN 20 134 INTERLEUKIN-5.  
FT DISULFID 63 63 INTERCHAIN (WITH C-105 OF OTHER CHAIN).  
FT FT (BY SIMILARITY).  
FT DISULFID 105 105 INTERCHAIN (WITH C-63 OF OTHER CHAIN).  
FT FT (BY SIMILARITY).  
FT CARBOHYD 76 76 POTENTIAL.  
FT CARBOHYD 90 90 POTENTIAL.  
FT CONFLICT 104 105 KC -> NF (IN REF. 2).  
FT CONFLICT 108 111 ERWR -> KKWK (IN REF. 2).  
FT CONFLICT 114 117 K -> N (IN REF. 2).  
FT CONFLICT 117 121 D -> N (IN REF. 2).  
FT CONFLICT 121 126 V -> F (IN REF. 2).  
FT CONFLICT 125 126 VI -> LL (IN REF. 2).  
SQ SEQUENCE 134 AA; 13224 MW; 87D18DB8F8CAC820 CRC64;  
Query Match 1.9%; Score 7; DB 1; Length 134;  
Best Local Similarity 100.0%; Pred. No. 2.24e+00; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 110 WRVKFL 116  
Qy 255 WRVKFL 261  
RESULT 9  
ID YIEH\_ECOLI STANDARD; PRT; 221 AA.  
AC P31467;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHEICAL 24.7 KD PROTEIN IN TNAB-BGLE INTERGENIC REGION.  
GN YIEH.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
RN [1]  
RC SEQUENCE FROM N.A.  
RX STRAIN-K12 / MG1655;  
RX MEDLINE; 93315143.  
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;  
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli  
genome: organizational symmetry around the origin of replication.;  
RL Genomics 16:551-561(1993).  
CC -!- SIMILARITY: BELONGS TO THE CBBY/CBBZ/GPH/YIEH FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC EMBL; L10328; AAA62066.1; -;  
CC EMBL; AE000448; AAC76738.1; -;  
CC ECOGENE; EG11725; YIEH.  
CC PRAM; PF00702; Hydrolase; 1.  
KW Hypothetical protein  
SQ SEQUENCE 221 AA; 24664 MW; C50CF0A788D7CE5F CRC64;  
Query Match 1.9%; Score 7; DB 1; Length 221;  
Best Local Similarity 100.0%; Pred. No. 2.24e+00; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 66 LAKEAE 72  
Qy 316 LAKEAE 322  
RESULT 10  
ID COPD\_PSESM STANDARD; PRT; 310 AA.  
AC P12377;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT COPPER RESISTANCE PROTEIN D.  
GN COPD  
OS Pseudomonas syringae (pv. tomato).  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;  
OC Pseudomonas.  
RN [1]  
RX MEDLINE; 88227880.  
RA Mellano M.A., Cooksey D.A.;  
RT "Nucleotide sequence and organization of copper resistance genes from  
Pseudomonas syringae pv. tomato.;  
RL J. Bacteriol. 170:2879-2883(1988).  
RN [2]  
RP SEQUENCE OF 278-310 FROM N.A.  
RA Mills S.D., Jasalavich C.A., Cooksey D.A.;  
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: EXACT FUNCTION NOT KNOWN. INVOLVED IN COPPER RESISTANCE.  
CC -!- APPEARS TO BE INVOLVED IN COPPER UPTAKE IN CONJUNCTION WITH COCP.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE  
(POTENTIAL).  
CC -!- INDUCTION: COPPER.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC EMBL; M19930; AAA25809.1; -;  
CC EMBL; L05176; AAA25802.1; -;  
CC PIR; D32018; D32018.  
KW Copper; Transmembrane; Inner membrane; Plasmid.  
FT TRANSMEM 11 31 POTENTIAL.  
FT TRANSMEM 47 67 POTENTIAL.  
FT TRANSMEM 92 112 POTENTIAL.  
FT TRANSMEM 116 136 POTENTIAL.  
FT TRANSMEM 157 177 POTENTIAL.  
FT TRANSMEM 201 221 POTENTIAL.  
FT TRANSMEM 233 253 POTENTIAL.  
FT TRANSMEM 282 302 POTENTIAL.  
SQ SEQUENCE 310 AA; 33089 MW; 6B47AEDA189B0E4 CRC64;  
Query Match 1.9%; Score 7; DB 1; Length 310;  
Best Local Similarity 100.0%; Pred. No. 2.24e+00; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 121 LLMVSL 127  
Qy 241 LLMVSL 247  
RESULT 11  
ID SRG7\_CAEEL STANDARD; PRT; 339 AA.  
AC P54129;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE SRG-7 PROTEIN.  
GN SRG-7 OR C18F10.8.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL NZ;  
RA Waterston R.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).

Match 1.9%; Score 7; DB 1; Length 357;

**1. Introduction**

RELATED PROTEIN).

NEEDED PROTEIN):

NEEDED PROTEIN):

US-09-376-430-2.rsp

Thu May 11 06:50:39 2000

GN DPEP1.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 17-40.  
RC TISSUE=KIDNEY;  
RX MEDLINE; 92074997.  
RA Igarashi P., Karniski L.P.;  
RT "Cloning of cDNAs encoding a rabbit renal brush border membrane  
RT protein immunologically related to band 3. Sequence similarity with  
RT microsomal dipeptidase.";  
RL Biochem. J. 280:71-78(1991).  
CC -!- FUNCTION: HYDROLYZES A WIDE RANGE OF DIPEPTIDES. IMPLICATED IN THE  
CC RENAL METABOLISM OF GLUTATHIONE AND ITS CONJUGATES. CONVERTS  
CC LEUKOTRIENE D4 TO LEUKOTRIENE E4; IT MAY PLAY AN IMPORTANT ROLE IN  
CC THE REGULATION OF LEUKOTRIENE ACTIVITY.  
CC -!- CATALYTIC ACTIVITY: DIPEPTIDE + H(2)O = 2 AMINO ACID.  
CC -!- COFACTOR: ZINC.  
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR;  
CC BRUSH BORDER MEMBRANE.  
CC -!- PTM: THE PRECISE POSITION OF THE C-TERMINUS AND GPI-ANCHOR OF THE  
CC MATURE RENAL DIPEPTIDASE IS NOT YET KNOWN.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M19.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

EMBL; X61503; CAA43720.1; -;  
PIR; S18442; S18442.  
DR PROSITE; PS00869; RENAL DIPEPTIDASE; 1.  
DR PFAM; PF01244; Renal dipeptase; 1.  
KW Hydrolase; Dipeptidase; Microsome; Signal; GPI-anchor; Glycoprotein;  
KW Zinc  
ZINC SIGNAL 1 16  
CHAIN 17 384 MICROSOMAL DIPEPTIDASE.  
FT CHAIN 17 384 REMOVED IN MATURE FORM (BY SIMILARITY).  
FT PROPEP 385 410 BY SIMILARITY.  
FT ACT\_SITE 141 141 ZINC (CATALYTIC) (POTENTIAL).  
FT METAL 286 286 POTENTIAL.  
FT CARBOHYD 57 57 GPI-ANCHOR (BY SIMILARITY).  
FT LIPID 384 384 2033F7BE9BF2CC72 CRC64;  
SQ SEQUENCE 410 AA; 45304 MW; 2033F7BE9BF2CC72 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 410;  
Best Local Similarity 100.0%; Pred. No. 2.24e+00;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 401 SLLLSL 407  
QY 245 SLLLSL 251

RESULT 15  
ID THCB\_RHOER STANDARD; PRT; 436 AA.  
AC P43492;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-FEB-2000 (Rel. 39, Last annotation update)  
DE CYTOCHROME P450 116 (EC 1.14.-.-).  
GN THCB OR CYP116.  
OS Rhodococcus erythropolis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.  
RC STRAIN=NI86/21.  
RX MEDLINE; 95138028.

RA Nagy I., Schoofs G., Compennolle F., Proost P., Vanderleyden J.,  
de Mot R.;  
RT "Degradation of the thiocarbamate herbicide EPTC (S-ethyl  
RT dipropylcarbamothioate) and biosafening by Rhodococcus sp. strain  
RT NI86/21 involve an inducible cytochrome P-450 system and aldehyde  
RT dehydrogenase.";  
RL J. Bacteriol. 177:676-687(1995).  
CC -!- FUNCTION: DEGRADATION OF THIOCARBAMATE HERBICIDES.  
CC -!- INDUCTION: BY EPTC (S-ETHYL DIPROPYL CARBAMOTHIOATE).  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

EMBL; U17130; RAC45749.1; -;  
DR HSRP; P33006; ICPT  
DR PROSITE; PS00869; CYTOCHROME\_P450; 1.  
DR PFAM; PF00067; P450; 1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Heme.  
FT INIT\_MET 0  
FT BINDING 374 374 HEME (BY SIMILARITY).  
SQ SEQUENCE 436 AA; 48796 MW; E24872BB47EESAS5 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 436;  
Best Local Similarity 100.0%; Pred. No. 2.24e+00;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 158 GDLLEYV 164  
QY 144 GDLLEYV 150

Search completed: Wed May 10 14:33:51 2000  
Job time : 109 secs.



\*\*\*\*\*  
[W][P][S][R][L][H]  
\*\*\*\*\*  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd  
MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 14:34:08 2000; MasPar time 256.31 Seconds  
Tabular output not generated.  
100.361 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (1-371) from US09376430A.pap  
Sequence: 1 MGRVLVWGAFLGGWMA.....DVTIGGFTFYNDRSYVAL 371  
Scoring table: TABLE uniprotatable  
Gap 60  
Searched: 225878 seqs, 69334122 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: sptrembl12  
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-nhc 8:sp-organelle  
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus  
Statistics: Mean 3.626; Variance 0.444; scale 8.171

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match	Length DB	ID	Description	Pred. No.	
1	8	2.2	22	6 Q28683	CYTCHROME P-450BC2 DN	6.00e-02	
2	8	2.2	56	6 Q00172	CYTCHROME P450 IIC1 (	6.00e-02	
3	8	2.2	216	5 Q9Y042	EGL-17.	6.00e-02	
4	8	2.2	271	11 Q35424	IABETA2 SUBUNIT.	6.00e-02	
5	8	2.2	281	7 Q31143	HISTOCOMPATIBILITY 2,	6.00e-02	
6	8	2.2	282	10 Q65009	BURP DOMAIN CONTAINING	6.00e-02	
7	8	2.2	282	12 Q9PXR2	OLIGOPEPTIDE ABC TRANS	6.00e-02	
8	8	2.2	660	4 Q15550	ALTERNATIVELY SPLICED	5.16e+00	
9	8	2.2	66	4 Q9Y4X3	CCL27 CHEMOKINE.	5.16e+00	
10	7	1.9	112	2 Q9X1T5	HYPOPHETICAL 14.0 KD P	5.16e+00	
11	7	1.9	120	2 Q60542	PERSEPHIN.	5.16e+00	
12	7	1.9	156	4 Q95W2	ORTHO-HALOGENOATE 1.2	5.16e+00	
13	7	1.9	176	10 Q38501	CYTCHROMIC CYCLOPHILIN	5.16e+00	
14	7	1.9	212	2 Q45498	HYPOPHETICAL 24.6 KD P	5.16e+00	
15	7	1.9	216	9 Q80097	CBF 59.	5.16e+00	
16	7	1.9	284	1 Q58324	284AA LONG HYPOPHETICA	5.16e+00	
17	7	1.9	287	2 P96167	POTATIVE ALDOLASE.	5.16e+00	
18	7	1.9	316	14 Q65566	CAPSID PROTEIN.	5.16e+00	
19	7	1.9	316	4 Q9Y400	HYPOPHETICAL 35.7 KD P	5.16e+00	
20	7	1.9	323	4 Q13624	INTERLEUKIN ENHANCER B	5.16e+00	

21	7	1.9	331	4 Q75752	GALT4 PROTEIN.	5.16e+00	
22	7	1.9	331	11 Q54906	UDP-GAL:BETAGLUCNAC BET	5.16e+00	
23	7	1.9	344	8 Q92TU2	NADH DEHYDROGENASE SUB	5.16e+00	
24	7	1.9	344	8 Q79890	NADH DEHYDROGENASE SUB	5.16e+00	
25	7	1.9	344	8 Q79893	NADH DEHYDROGENASE SUB	5.16e+00	
26	7	1.9	365	4 Q60509	NSP-LIKE 1 (RTN2-B) (F	5.16e+00	
27	7	1.9	376	5 Q20934	SIMILAR TO 4-HYDROXYBE	5.16e+00	
28	7	1.9	376	10 Q39353	CELL WALL-PLASMA MEMBER	5.16e+00	
29	7	1.9	399	5 Q9XWT8	Y102FSA.1 PROTEIN	5.16e+00	
30	7	1.9	437	10 Q48902	MALATE DEHYDROGENASE (	5.16e+00	
31	7	1.9	487	4 Q16756	LIVER CYTOCHROME P450	5.16e+00	
32	7	1.9	506	2 Q53835	PHASE-2 FLAGELLIN STRU	5.16e+00	
33	7	1.9	506	2 Q53837	PHASE-2 FLAGELLIN STRU	5.16e+00	
34	7	1.9	516	4 Q14682	PIG10.	5.16e+00	
35	7	1.9	544	5 Q02081	SIMILARITY TO A SHORT	5.16e+00	
36	7	1.9	589	4 Q75464	NUCLEAR MATRIX PROTEIN	5.16e+00	
37	7	1.9	609	4 Q13623	INTERLEUKIN ENHANCER B	5.16e+00	
38	7	1.9	642	10 Q22041	NP1-RELATED PROTEIN K	5.16e+00	
39	7	1.9	645	10 Q92T08	RECEPTOR-LIKE PROTEIN	5.16e+00	
40	7	1.9	746	5 Q21843	SIMILAR TO DROSOPHILIA	5.16e+00	
41	7	1.9	751	2 Q9WZ18	(P)PPGPP SYNTHETASE.	5.16e+00	
42	7	1.9	771	11 Q97691	INS-1 WINGED HELIX.	5.16e+00	
43	7	1.9	1685	10 Q04142	RNA POLYMERASE II LARG	5.16e+00	
44	7	1.9	2401	5 Q6216	RHOPTRY PROTEIN (FRAGM	5.16e+00	
45	7	1.9	2895	10 Q65551	HYPOTHETICAL 326.6 KD	5.16e+00	

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	22 AA.
ID	Q28683			
AC	Q28683			
DT	01-NOV-1996 (TREMELrel. 01, Created)			
DT	01-NOV-1996 (TREMELrel. 01, Last sequence update)			
DE	CYTCHROME P-450BC2 DNA, 5' FLANKING REGION (FRAGMENT).			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
RA	SEQUENCE FROM N.A.			
RA	MEDLINE: 87053173.			
RA	GOVIND S., BELL P.A., KEMPER B.;			
RT	"Structure of genes in the cytochrome p-450Bc subfamily: conservation			
RT	of intron locations in the phenobarbital-inducible family.";			
RL	DNA 3:371-382(1586).			
DR	EMBL: M14955; AAA31208.1; -			
FT	NON-TER 22 22			
SQ	SEQUENCE 22 AA; 2460 MW; BCB03755 CRC32;			

Query Match 2.2%; Score 8; DB 6; Length 22;  
Best Local Similarity 100.0%; Pred. No. 6.00e-02;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	14 LLLSLWK 21			
QY	246 LLLSLWK 253			
RESULT	2	PRELIMINARY;	PRT;	56 AA.
ID	Q00172			
AC	Q00172			
DT	01-NOV-1996 (TREMELrel. 01, Created)			
DT	01-NOV-1996 (TREMELrel. 01, Last sequence update)			
DE	CYTCHROME P450 IIC1 (EC 1.4.14.1) (PHENOBARBITAL-INDUCIBLE)			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
RA	SEQUENCE FROM N.A.			

RC STRAIN-NEW ZEALAND WHITE; TISSUE-LIVER;  
 RX MEDLINE; 90197893.  
 RA ZHAO J., CHAN G., GOVIND S., BELL P., KEMPER B.W.;  
 RT "Structure of 5' regions and expression of phenobarbital-inducible  
 RL DNA Cell Biol. 9:37-48(1990).  
 DR EMBL; M74199; AAA31436.1; -;  
 FT NON\_TER 56  
 SQ SEQUENCE 56 AA; 5968 MW; D6432166 CRC32;

Query Match 2.2%; Score 8; DB 6; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 6.00e-02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 14 LLLLSLWK 21  
 QY 246 LLLLSLWK 253

RESULT 3  
 ID O9Y042 PRELIMINARY; PRT; 216 AA.  
 AC O9Y042;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE EGL-17.  
 GN EGL-17.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-N2;  
 RA BURDINE R.D., CHEN E.B., KWOK S.F., STERN M.J.;  
 RT "egl-17 encodes an invertebrate fibroblast growth factor family member  
 RT required specifically for sex myoblast migration in Caenorhabditis  
 RL elegans.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(1997).  
 DR EMBL; U85766; AAD00574.1; -;  
 SQ SEQUENCE 216 AA; 25224 MW; D26A23FD CRC32;

Query Match 2.2%; Score 8; DB 5; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 6.00e-02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 171 TPRPKLS 178  
 QY 223 TPRPKLS 230

RESULT 4  
 ID O35424 PRELIMINARY; PRT; 271 AA.  
 AC O35424;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DE IABETA2 SUBUNIT.  
 GN H2-IABETA2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ROMEN L., QIN S., AHEARN M.E., LORETZ C., FAUST J., LASKY S.,  
 RA MAHAIRAS G., HOOD L.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF027865; AAB81530.1; -;  
 DR PFAM; PF00047; 1g; 1.  
 DR PFAM; PF00969; MHC\_II\_beta; 1.  
 SQ SEQUENCE 271 AA; 30464 MW; 5AD2E2D1 CRC32;

Query Match 2.2%; Score 8; DB 11; Length 271;  
 Best Local Similarity 100.0%; Pred. No. 6.00e-02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 228 GAAVFLG 235  
 QY 9 GAAVFLG 16

RESULT 5  
 ID O31143 PRELIMINARY; PRT; 281 AA.  
 AC O31143;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE HISTOCOMPATIBILITY 2, O REGION BETA LOCUS (MHC CLASS II A-BETA-2).  
 GN H2-OB.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SPLEEN;  
 RX LARHAMMAR D., HAMMERLING U., RASK L., PETERSON P.A.;  
 RA "Sequence of gene and cDNA encoding murine major histocompatibility  
 RT complex class II gene A beta 2.";  
 RL J. Biol. Chem. 260:14111-14119(1985).  
 DR EMBL; M19423; AAA51637.1; -;  
 DR EMBL; M11800; AAA51637.1; -;  
 DR MGD; MGI:95925; H2-Ob.  
 DR PFAM; PF00047; 1g; 1.  
 DR PFAM; PF00969; MHC\_II\_beta; 1.  
 KW Membrane; MHC.  
 SQ SEQUENCE 281 AA; 31642 MW; FA13A147 CRC32;

Query Match 2.2%; Score 8; DB 7; Length 281;  
 Best Local Similarity 100.0%; Pred. No. 6.00e-02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 228 GAAVFLG 235  
 QY 9 GAAVFLG 16

RESULT 6  
 ID O65009 PRELIMINARY; PRT; 282 AA.  
 AC O65009;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DE BURP DOMAIN CONTAINING PROTEIN.  
 OS Brassica napus (Rape).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Brassica.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. TOPAS;  
 RA HATTORI J., BOUTILLIER K., VAN LOOKEREN CAMPAGNE M., MIKI B.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF049028; AAC15700.1; -;  
 SQ SEQUENCE 282 AA; 32089 MW; 38165A7C CRC32;

Query Match 2.2%; Score 8; DB 10; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 6.00e-02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 15 SLLLSLW 22  
 QY 245 SLLLSLW 252

RESULT 7

\*\*\*\*\*  
 W P E R L E H  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MParch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 11:09:16 2000; MasPar time 24.54 Seconds  
 713.186 Million cell updates/sec  
 Tabular output not generated.

Title: >US-09-376-430-2  
 Description: (1-371) from US09376430A.pep (1 of 25)  
 Perfect Score: 2788  
 Sequence: 1 MGRLLWGAAVFLGGWMA.....DVTIGGTFVMDRYSVAL 371

Scoring table: PAM 150  
 Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: pir62  
 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 48.184; Variance 95.611; scale 0.504

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	210	7.5	373	2	A55718	interleukin-2 recepto
2	195	7.0	369	2	A42565	interleukin-2 recepto
3	188	6.7	369	2	I49280	interleukin-2 recepto
4	178	6.4	467	2	I56896	gene gfi-2 protein
5	176	6.3	522	2	B45268	interleukin-9 recepto
6	166	6.0	468	2	A45268	interleukin-9 recepto
7	164	5.9	507	1	A46713	erythropoietin recept
8	163	5.8	507	1	A32385	erythropoietin recept
9	160	5.7	508	1	ZUHUR	erythropoietin recept
10	115	4.1	878	1	A40091	interleukin-3 recepto
11	113	4.1	897	1	A39255	cytokine receptor com
12	110	3.9	326	2	H71497	probable yop transloc
13	109	3.9	430	2	E71350	probable aspartate am
14	108	3.9	490	2	B28516	cytochrome P450 2C7
15	110	3.9	551	2	A30342	interleukin-2 recepto
16	110	3.9	634	2	S33339	somatotropin receptor
17	110	3.9	638	2	S12136	somatotropin receptor
18	110	3.9	638	2	S04530	somatotropin receptor
19	110	3.9	638	2	A35991	somatotropin receptor
20	108	3.9	896	1	A35782	cytokine receptor com
21	108	3.8	896	1	S40149	integrin alpha-7C cha
22	107	3.8	298	2	B34791	interleukin-7 recepto
23	105	3.8	378	2	S00842	leukosialin precursor

24	107	3.8	459	2	A34791	interleukin-7 recepto
25	106	3.8	490	2	A36122	cytochrome P450 2C13
26	106	3.8	1073	1	OYHUX	heat-stable enterotox
27	106	3.8	1106	2	S38783	integrin alpha chain
28	105	3.8	1135	2	I61186	alpha-7 integrin - mo
29	102	3.7	386	1	QOBES0	latent membrane prote
30	104	3.7	430	2	H72604	hypothetical protein
31	102	3.7	443	2	G75038	probable na+/h+ anti
32	102	3.7	861	2	S77086	hypothetical protein
33	102	3.7	1062	2	JC5951	integrin alpha 7 chai
34	102	3.7	1137	2	JC5950	integrin alpha 7 chai
35	104	3.7	1151	2	A45226	integrin alpha-1 chai
36	104	3.7	1180	2	A35854	integrin alpha-1 chai
37	100	3.6	169	2	S76289	hypothetical protein
38	101	3.6	184	2	H72248	ribosomal protein L5
39	100	3.6	220	2	I80329	receptor tyrosine kin
40	101	3.6	359	2	I49341	MIP-1 alpha receptor
41	101	3.6	381	2	S24611	latent membrane prote
42	101	3.6	404	1	IABECA	latent membrane prote
43	101	3.6	418	2	D70038	maltoextrin transpor
44	100	3.6	876	2	A49508	protein-tyrosine kina
45	100	3.6	1021	2	I39207	leukocyte surface pro

## ALIGNMENTS

RESULT 1  
 ENTRY  
 TITLE #type complete  
 ORGANISM interleukin-2 receptor gamma chain precursor - dog  
 DATE 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 23-Jul-1999  
 A55718  
 #accessions A55718  
 #references A55718  
 #authors Heathorn, P.S.; Somberg, R.L.; Fimiani, V.M.; Puck, J.M.;  
 Patterson, D.F.; Pelsburg, P.J.  
 #journal Genomics (1994) 23:593-74  
 #title IL-2Rgamma gene microdeletion demonstrates that canine X-linked severe combined immunodeficiency is a homologue of the human disease.  
 #cross-references MUID:95130114  
 #accession A55718  
 #status preliminary  
 #molecule\_type mRNA  
 #residues 1-373 #label HEN  
 #cross-references GB:U04361; NID:G517411; PIDN:AAC48403.1; PID:G517412  
 #superfamily interleukin-2 receptor gamma chain  
 #keywords cytokine receptor; duplication  
 #summary #length 373 #molecular-weight 42516 #checksum 7960

Query Match	7.5%	Score 210;	DB 2;	Length 373;
Best Local Similarity	25.9%	Pred. No. 4,31e-17;		
Matches	49;	Conservative	56;	Mismatches 73;
				Indels 11; Gaps 10;
Db	59	EVQCFVNVFMNCTWSSSEPRPTNLTWHYKNSNDKVOEGCHYLFSEVITAGCWLQ	118	
Qy	31	QIQIYIENLEFVQVWNAKYSR-TNLTFFHYRF-NGDEA-YDQCTNYLLQEGHSGCLLD	87	
Db	119	KEEHLXETFFVQLDRPREPRQSTQKLQNLVWPAPENLTNLHNSOLEWSNRH	178	
Qy	88	AEQRD--DILYFSIRNGTHPVFTAS-RWVYYIL-KPSSPKHVR-FSWHQDAVITCSDLS	142	
Db	179	LDHCLIEHVQVRSVDWRSWTQSQVDHNSFSLPSVDGQKFTFVRVRSY-NPLCGSAQRW	237	
Qy	143	YGD-LLYEVQVRSFPDTEW-QSKQENTCNVTIEGLDAEKYSFWVRVAKMEDVYGPDTYP	200	
Db	238	SEWSHPHWH 246		
Qy	201	SDWSEVTCW 209		

RESULT 2  
 ENTRY  
 TITLE #type complete  
 A42565

```

TITLE      interleukin-2 receptor gamma chain - human
ORGANISM   #formal_name Homo sapiens #common_name man
DATE       04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change
REFERENCE  23-Jul-1999
AUTHORS     Takeshita, T.; Asao, H.; Ohtani, K.; Ishii, N.; Kumaki, S.;
            Tanaka, N.; Munakata, H.; Nakamura, M.; Sugamura, K.
JOURNAL     Science (1992) 257:379-382
#journal    Cloning of the gamma chain of the human IL-2 receptor.
#title      Cross-references MUID:92335883
#accession  A42565
#status     preliminary; not compared with conceptual translation
#molecule_type nucleic acid; protein
#residues    1-369 #label TAK
#cross-references GB:D1086; NID:g303611; PIDN:BAA01857.1;
            PID:g1002334; PID:g219890
#experimental_source MOLT beta lymphoid cells
#note       sequence extracted from NCBI backbone (NCBIP:109167)
REFERENCE   A46591
AUTHORS     Noguchi, M.; Adelstein, S.; Cao, X.; Leonard, W.J.
JOURNAL     J. Biol. Chem. (1993) 268:13601-13608
#journal    Characterization of the human interleukin-2 receptor gamma
#title      chain gene.
#cross-references MUID:93293887
#accession  A46591
#status     preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues    1-369 #label RES
#cross-references GB:L12183; NID:g307056; PIDN:AAA59145.1; PID:g307058
REFERENCE   I54332
AUTHORS     Puck, J.M.; Deschenes, S.M.; Porter, J.C.; Dutra, A.S.;
            Brown, C.J.; Willard, H.F.; Henthorn, P.S.
JOURNAL     Hum. Mol. Genet. (1993) 2:1099-1104
#journal    The interleukin-2 receptor gamma chain maps to Xq13.1 and is
#title      mutated in X-linked severe combined immunodeficiency,
            SCIDX1
#cross-references MUID:94004847
#accession  I54332
#status     preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues    1-369 #label RE2
#cross-references GB:L19546; NID:g349631; PIDN:AAC37524.1; PID:g349632
GENETICS    #gene      GDB:IL2RG; SCIDX1; IMD4
            #cross-references GDB:I34807; OMIM:308380
            #map_position Xq13.1-Xq13.1
            #introns    39/1; 90/2; 152/1; 198/3; 253/1; 285/2; 308/3
            #note      defects are associated with an X-linked form of severe
            combined immunodeficiency
CLASSIFICATION #superfamily interleukin-2 receptor gamma chain
KEYWORDS        cytokine receptor; duplication; immunodeficiency; severe
            combined immunodeficiency; transmembrane protein
SUMMARY         #length 369 #molecular_weight 42287 #checksum 7850
Query Match    7.08; Score 195; DB 2; Length 369;
Best Local Similarity 25.98; Pred. No. 1.51e-14;
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;
Db 59 EVQCFVFNVEYMNCTWNSSEPOPTNLTLYWYKNSDNKQKCSHYLFSEETSGCOLQ 118
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
31 QIQIIVFNLETVQVTWNASYSR-TNLTFRYF-NGD-EAYDQCTNVLQEGHTSGCLLD 87
Db 119 KKEIHLXQIFVQLOQDPREPRRATOMLQNLVWAPENLTLHKUSESOLELNNWNR 178
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
88 AEQRDILYFSIR-NGTH-PVFSTRWM-VYIL-KPSSPKHVRP-SWHQDAVTTCSDLS 142
Db 179 LNHCLHLYQRYRDWDHSTEQSDVYRHKFSLPSVDQKTYTFVRSR-FNPLGCSAQHW 237
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
143 YGD-LLYEYQVRSFPTETW-QSQKQNTCNVTIEGLDAEKCYSPWVRKAMEDVYGPDP 200
Db 238 SEMSHPIHW 246
            : : : :

```

```

QY 201 SDWSEVTCW 209
RESULT      3
ENTRY       I49280 #type complete
TITLE       interleukin-2 receptor gamma chain precursor - mouse
ORGANISM    #formal_name Mus musculus #common_name house mouse
DATE        27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change
            23-Jul-1999
ACCESSIONS  I49280; A47514; JN0592; JN0775; S37582; I53398
REFERENCE    A47514
AUTHORS     Cao, X.; Kozak, C.A.; Liu, Y.
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. (1993) 90:8464-8468
#journal    Characterization of cDNAs encoding the murine interleukin 2
#title      receptor (IL-2R) gamma chain: Chromosomal mapping and
            tissue specificity of IL-2R gamma chain expression.
#cross-references MUID:93391374
#accession  I49280
#status     preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues    1-369 #label CAO
#cross-references EMBL:U21795; NID:g727349; PIDN:AAA64279.1;
            PID:g727350
#accession  A47514
#status     translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues    1-369 #label RE2
#cross-references GB:L20048; NID:g404067; PIDN:AAA39286.1; PID:g404068
REFERENCE    JN0592
AUTHORS     Kumaki, S.; Kondo, M.; Takeshita, T.; Asao, H.; Nakamura, M.;
            Sugamura, K.
JOURNAL     Biochem. Biophys. Res. Commun. (1993) 193:356-363
#journal    Cloning of the mouse interleukin 2 receptor gamma chain:
#title      Demonstration of functional differences between the mouse
            and human receptors.
#cross-references MUID:93277575
#accession  JN0592
#status     nucleic acid sequence not shown
#molecule_type mRNA
#residues    1-369 #label KDM
#cross-references DDBJ:D13565; NID:g303684; PIDN:BAA02760.1;
            PID:d1003265; PID:g303685
REFERENCE    JN0775
AUTHORS     Kobayashi, N.; Nakagawa, S.; Minami, Y.; Taniguchi, T.; Kono,
            T.
JOURNAL     Gene (1993) 130:303-304
#journal    Cloning and sequencing of the cDNA encoding a mouse IL-2
#title      receptor gamma.
#cross-references MUID:93366191
#accession  JN0775
#molecule_type mRNA
#residues    1-369 #label KOB
#cross-references GB:D13821; NID:g436045; PIDN:BAA02974.1;
            PID:d1003480; PID:g436046
REFERENCE    S37582
AUTHORS     Chiu, R.K.; Dougherty, G.J.
JOURNAL     Submitted to the EMBL Data Library, October 1993
#journal    Regulation of CD44-mediated cellular adhesion by the IL-2 R
#title      gamma chain.
#cross-references MUID:93375337
#accession  S37582
#status     preliminary
#molecule_type mRNA
#residues    1-350, 'S', 352-366, 'S', 368-369 #label CHI
#cross-references EMBL:X75337
REFERENCE    I53398
AUTHORS     DiSanto, J.P.; Certain, S.; Wilson, A.; MacDonald, H.R.;
            Ayner, P.; Fischer, A.; de Saint Basile, G.
JOURNAL     Eur. J. Immunol. (1994) 24:3014-3018
#journal    The murine interleukin-2 receptor gamma chain gene:
#title      organization, chromosomal localization and expression in
            the adult thymus.
#cross-references MUID:95104285
#accession  I53398

```



```

##status      Preliminary; translated from GB/EMBL/DBDJ
##molecule_type DNA
##residues    1-369 #label RES
##cross-references GB:575852; NID:9861554; PIDN:AAB32904.1; PID:9861555
GENETICS
#gene         IL-2Rgamma
#introns      39/1; 90/2; 152/1; 199/3; 254/1; 286/2; 308/3
#complex      The high affinity receptor is a heterotrimer of alpha (see
               PIR:UHS2), beta (see PIR:A35052), and gamma chains;
               heterodimers of alpha or beta and gamma chains are
               intermediate affinity receptors.
FUNCTION
#description  receptor for interleukin-2
#pathway     Interleukin-2 stimulated growth and differentiation of T
               cells, B cells, NK cells, LAK cells, monocytes,
               macrophages, and oligodendrocytes
CLASSIFICATION #superfamily interleukin-2 receptor gamma chain
KEYWORDS       cytokine receptor; duplication; glycoprotein; transmembrane
               protein
FEATURE
1-22          #domain signal sequence #status predicted #label SIG\
23-369        #product interleukin-2 receptor gamma chain #status
               predicted #label MAT\
256-284       #domain transmembrane #status predicted #label TM\
71,75,84,96,159, #binding_site carbohydrate (Asn) (covalent) #status
184,306        predicted
SUMMARY        #length 369 #molecular-weight 42241 #checksum 6734
               Predicted
Query Match    6.7%; Score 188; DB 2; Length 369;
Best Local Similarity 26.8%; Pred. No. 2.21e-13; Indels 12; Gaps 11;
Matches 51; Conservative 52; Mismatches 75;
Db 59 EQQCFVNIENMNCNSSEPATNLTHRYKYVSDNNNTFQCSHYLSKEITSCQIQ 118
QY 31 QIQIIVNLETQVTVNASKYSR-TNLTHRYFN-GDE-AYDQCTNLLQEGHTSGCLLD 87
Db 119 KEDIQIYTFVQLODPQRPORRAVKLNLQNLVIRAPENLTLSLSQELRWKSRH 178
QY 88 AEQRDILYFSIR-NGTH-PVFTASRWV-VYIL-KPSSPKHVRFS-WHQDAVTV--TCSD 140
Db 179 IKERCQYLVOYRNSRDRSNTLIVNHPFSLPSVDELKRYFRVRSRY-NPICGSSQ 237
QY 141 LSYGDLLEYQVRSPPDTEW-QSKQENTCNVTIEGLDAEKYCFWVRVKAMEDVYGPDTY 199
Db 238 WSKWSQPVHW 247
QY 200 PSDWSEVTCW 209
RESULT 4
ENTRY     I56896 #type complete
TITLE     gene gfi-2 protein - rat
ORGANISM  #formal_name Rattus norvegicus #common_name Norway rat
DATE      26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
          28-Feb-1997
ACCESSIONS
REFERENCE  I56896
AUTHORS    Flubacher, M.M.; Bear, S.E.; Tschlis, P.N.
JOURNAL    J. Virol. (1994) 68:7709-7716
#title     Replacement of interleukin-2 (IL-2)-generated mitogenic
           signals by a myeloid cell focus-forming (MCF) or xenotropic
           virus-induced IL-9-dependent autocrine loop: implications
           for MCF virus-induced leukemogenesis.
#cross-references MUID:9505595
#accession  I56896
#status     Preliminary; translated from GB/EMBL/DBDJ
#molecule_type mRNA
#residues   1-467 #label RES
#cross-references GB:L36459; NID:9598371; PID:9598372
GENETICS
#gene       gfi-2
SUMMARY     #length 467 #molecular-weight 52017 #checksum 3280

```

```

Query Match    6.4%; Score 178; DB 2; Length 467;
Best Local Similarity 30.0%; Pred. No. 9.74e-12;
Matches 24; Conservative 24; Mismatches 29; Indels 3; Gaps 3;
Db 272 ILVAVPIFLLTGLIHF-LFRLSPKVRKIFQYQNPSPFAFHFPLYSVYHGFQFTWIGARR 330
QY 233 ILISSLAILLMVSLLLSLKWL-RVKKFLIPSPVDPKSIFFGLFEIHQGNFQEWITDTQ 291
Db 331 AGPOARQ-DGASAPSGDSES 349
QY 292 NVHLHMKAGAEQESGPEEP 311
RESULT 5
ENTRY     B45268 #type complete
TITLE     Interleukin-9 receptor precursor - human
ORGANISM  #formal_name Homo sapiens #common_name man
DATE      27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change
          10-Sep-1997
ACCESSIONS
REFERENCE  B45268
AUTHORS    Renauld, J.C.; Druetz, C.; Kermouni, A.; Houssiau, F.;
           Uytendhove, C.; Van Roost, E.; Van Snick, J.
           Proc. Natl. Acad. Sci. U.S.A. (1992) 89:5690-5694
#journal   Expression cloning of the murine and human interleukin 9
#title     receptor cDNAs.
#cross-references MUID:92302307
#accession B45268
#status     Preliminary
#molecule_type mRNA
#residues   1-522 #label REN
#cross-references GB:M84747; NID:9184508; PID:9184509
KEYWORDS    glycoprotein; receptor; T-cell proliferation; transmembrane
           protein
SUMMARY     #length 522 #molecular-weight 57333 #checksum 2693
Query Match    6.3%; Score 176; DB 2; Length 522;
Best Local Similarity 27.8%; Pred. No. 2.06e-11;
Matches 50; Conservative 43; Mismatches 74; Indels 14; Gaps 12;
Db 224 LRVOMATLDDVVVEERYTGQSE---WSQPVCFQAPQRPQGLIPWGPNG-TLVAVSI 279
QY 183 FWVRVAMED-VYGPDTYPSDSEVTCWQGE-IRDCAEETP-TTPKPKLSKILISLLA 239
Db 280 FLLLTGPTYL-LFKLSPRVRKRIYQNPSPAMFFQPLYSVHNGNFQTMGAHAGVLLSQ 338
QY 240 ILLMVSLLLSLKWL-RVKKFLIPSPVDPKSIFFGLFEIHQGNFQEWITDTQNVHL-H 297
Db 339 DCAGTPQGALEPCVQETALLTCGPAPRWKSVALEEEQEGG-T-RLPGN-LSEEDVLPA 395
QY 298 KMAGAEQES-GPEELVVLQAKTEAPSRMLDPQTEKEASGSLQLPQLOGGDVVTI 356
Db 396 G 396
QY 357 G 357
RESULT 6
ENTRY     A45268 #type complete
TITLE     Interleukin-9 receptor precursor - mouse
ORGANISM  #formal_name Mus musculus #common_name house mouse
DATE      27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change
          10-Sep-1997
ACCESSIONS
REFERENCE  A45268
AUTHORS    Renauld, J.C.; Druetz, C.; Kermouni, A.; Houssiau, F.;
           Uytendhove, C.; Van Roost, E.; Van Snick, J.
           Proc. Natl. Acad. Sci. U.S.A. (1992) 89:5690-5694
#journal   Expression cloning of the murine and human interleukin 9
#title     receptor cDNAs.
#cross-references MUID:92302307
#accession  A45268

```

```

##status preliminary
##molecule_type mRNA
##residues 1-468 ##label REN
##cross-references GB:M84746; NID:g194044; PID:g194045
##note authors translated the codon GGG for residue 394 as Glu
KEYWORDS glycoprotein; receptor; T-cell proliferation; transmembrane
protein
SUMMARY #length 468 #molecular-weight 52260 #checksum 48

Query Match 6.0%; Score 166; DB 2; Length 468;
Best Local Similarity 24.7%; Pred. No. 8.32e-10;
Matches 20; Conservative 28; Mismatches 31; Indels 2; Gaps 2;

Db 273 ILVVVPFLLLTGTVHL-LFKLSRLKRIYQNTIPSEAFPHLYSVYHGDFQSWTGARR 331
QY 233 ILISLAILLMVSLLSLWKL-RVKFKLIPSPVDPKSIFFGLFEIHQGNFQEWITDQ 291
Db 332 AGPOARQNGVSTSSAGSESSI 352
QY 292 NVAHLHKMAGAEQESGPEPL 312

```

```

RESULT 7
ENTRY A46713 #type complete
TITLE erythropoietin receptor precursor - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change
ACCESSIONS A46713
REFERENCE A46713
#authors Masuda, S.; Nagao, M.; Takahata, K.; Konishi, Y.; Gallyas
Jr., F.; Tabira, T.; Sasaki, R.
#journal J. Biol. Chem. (1993) 268:11208-11216
#title Functional erythropoietin receptor of the cells with neural
characteristics. Comparison with receptor properties of
erythroid cells.
#cross-references MUID:93266574
#accession A46713
##status preliminary
##molecule_type mRNA
##residues 1-507 ##label MAS
##cross-references GB:D13566; NID:g286209; PIDN:BAA02761.1;
PID:d1003266; PID:g286210
##experimental_source PC12 and erythroid cells
##note sequence extracted from NCBI backbone (NCBIN:132811,
NCBIP:132813)
CLASSIFICATION #superfamily erythropoietin receptor; cytokine receptor
homology
KEYWORDS cytokine receptor; glycoprotein; transmembrane protein
FEATURE
1-24 #domain signal sequence #status predicted #label SIG\
25-507 #product erythropoietin receptor #status predicted
#label MAT\
25-249 #domain extracellular #status predicted #label EXT\
52-238 #domain cytokine receptor homology #label CRS\
250-271 #domain transmembrane #status predicted #label TM\
272-507 #domain intracellular #status predicted #label INT\
75 #binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY #length 507 #molecular-weight 55499 #checksum 7999

Query Match 5.9%; Score 164; DB 1; Length 507;
Best Local Similarity 45.2%; Pred. No. 1.73e-09;
Matches 28; Conservative 11; Mismatches 19; Indels 4; Gaps 3;

```

```

Db 247 LDPLILTLTL-LVLSLILLVALLSHRRRLQKIWPGIPSPSEFGLFTTHKGNFOL 305
QY 229 LSKFLLISSAILLMVSLLL--LSLWKLWRVKKFLI-PSVDPKSIFFGLFEIHQGNFQ 285
Db 306 WL 307
QY 286 WI 287

```

```

RESULT 8
ENTRY A32385 #type complete
TITLE erythropoietin receptor precursor, membrane-bound form -
mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 28-Sep-1990 #sequence_revision 05-Apr-1995 #text_change
22-Jun-1999
ACCESSIONS A41686; A32385; S13249
REFERENCE A41686
#authors Hino, M.; Tojo, A.; Misawa, Y.; Morii, H.; Takaku, F.;
Shibuya, M.
#journal Mol. Cell. Biol. (1991) 11:5527-5533
#title Unregulated expression of the erythropoietin receptor gene
caused by insertion of spleen focus-forming virus long
terminal repeat in a murine erythroleukemia cell line.
#cross-references MUID:92017832
#accession A41686
##molecule_type mRNA
##residues 1-507 ##label HIN
##cross-references GB:S59388; NID:g237036; PIDN:AAB20029.1; PID:g237037
##experimental_source murine erythroleukemia (MEL) cell line F5-5
REFERENCE A32385
#authors D'Andrea, A.D.; Lodish, H.F.; Wong, G.G.
#journal Cell (1989) 57:277-285
#title Expression cloning of the murine erythropoietin receptor.
#cross-references MUID:89195238
#accession A32385
##molecule_type mRNA
##residues 1-507 ##label DAA
##cross-references GB:J04843; NID:g193090; PIDN:AAA37571.1; PID:g309219
##experimental_source murine erythroleukemia (MEL) cells, subclone 745
REFERENCE S13249
#authors Kuramochi, S.; Ikawa, Y.; Todokoro, K.
#journal J. Mol. Biol. (1990) 216:567-575
#title Characterization of murine erythropoietin receptor genes.
#cross-references MUID:91080149
#accession S13249
##molecule_type DNA; mRNA
##residues 1-507 ##label KUR
##cross-references EMBL:X33081; NID:g50861; PIDN:CAA37248.1; PID:g50862
##experimental_source murine erythroleukemia K-1 cells
GENETICS 39/1; 83/2; 142/1; 194/3; 246/1; 275/2; 304/3
CLASSIFICATION #superfamily erythropoietin receptor; cytokine receptor
homology
KEYWORDS alternative splicing; cytokine receptor; glycoprotein;
transmembrane protein
FEATURE
1-24 #domain signal sequence #status predicted #label SIG\
25-507 #product erythropoietin receptor #status predicted
#label MAT\
25-249 #domain extracellular #status predicted #label EXT\
52-238 #domain cytokine receptor homology #label CRS\
250-271 #domain transmembrane #status predicted #label TM\
272-507 #domain intracellular #status predicted #label INT\
75 #disulfide_bonds #status predicted
#binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY #length 507 #molecular-weight 55194 #checksum 9070

Query Match 5.8%; Score 163; DB 1; Length 507;
Best Local Similarity 46.8%; Pred. No. 2.48e-09;
Matches 29; Conservative 9; Mismatches 20; Indels 4; Gaps 3;

```

```

Db 247 LDPLILTLTL-LVLSLILLVALLSHRRRLQKIWPGIPSPSEFGLFTTHKGNFOL 305
QY 229 LSKFLLISSAILLMVSLLL--LSLWKLWRVKKFLI-PSVDPKSIFFGLFEIHQGNFQ 285
Db 306 WL 307
QY 286 WI 287

```

```

9
RESULT 9
ENTRY ZUHUR #type complete
TITLE erythropoietin receptor precursor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 12-Feb-1993 #sequence_revision 03-Apr-1995 #text_change
22-Jun-1999

ACCESSIONS A43799; A60160; A49824; A53958; A52563
REFERENCE Jones, S.S.; D'Andr a, A.D.; Haines, L.L.; Wong, G.G.
#journal Blood (1990) 76:31-35
#title Human erythropoietin receptor: cloning, expression, and
Biologic characterization.
#cross-references MUID:90304340
#accession A43799
#molecule_type mRNA
#residues 1-508 #label JON
#cross-references GB:M60459; NID:g182244; PIDN:AAA52403.1; PID:g182245
REFERENCE A60160
#authors Winkelman, J.C.; Penny, L.A.; Deaven, L.L.; Forget, B.G.;
Jenkins, R.B.
#journal Blood (1990) 76:24-30
#title The gene for the human erythropoietin receptor: analysis of
the coding sequence and assignment to chromosome 19-p.
#cross-references MUID:90304334
#accession A60160
#status not compared with conceptual translation
#molecule_type mRNA; DNA
#residues 1-101; 'R', 103-188; 'RP', 191-243; 'E', 245-508 #label WIN
REFERENCE A49824
#authors Noguchi, C.T.; Bae, K.S.; Chin, K.; Wada, Y.; Schechter,
A.N.; Hankins, W.D.
#journal Blood (1991) 78:2548-2556
#title Cloning of the human erythropoietin receptor gene.
#cross-references MUID:92399733
#accession A49824
#molecule_type DNA
#residues 1-508 #label NOG
#cross-references GB:S45332; NID:g255496; PIDN:AAB23271.1; PID:g255497
#experimental_source placenta
#note sequence extracted from NCBI backbone (NCBIN:113293,
NCBIP:113294)
REFERENCE A53958
#authors Ehrenman, K.; St. John, T.
#journal Exp. Hematol. (1991) 19:973-977
#title The erythropoietin receptor gene: cloning and identification
of multiple transcripts in an erythroid cell line OCI-M1.
#cross-references MUID:91372359
#accession A53958
#molecule_type mRNA
#residues 1-508 #label EHR
REFERENCE A55280
#authors Penny, L.A.; Forget, B.G.
#journal Genomics (1991) 11:974-980
#title Genomic organization of the human erythropoietin receptor
gene.
#cross-references MUID:92147143
#accession A55280
#molecule_type DNA
#residues 1-17; 381-387; 'LLEQQODA', 391-395; 504-508 #label PEN
#note sequence modified after extraction from NCBI backbone
#note the authors translated the codon GAT for residue 31 to B
#note an insert compared to other published sequences is
considered by authors as likely cloning artifact,
rather than polymorphism
REFERENCE I52563
#authors Maouche, L.; Tournamille, C.; Hattab, C.; Boffa, G.; Carttron,
J.P.; Chretien, S.
#journal Blood (1991) 78:2557-2563
#title Cloning of the gene encoding the human erythropoietin
receptor.
#cross-references MUID:92399734
#accession I52563

```

```

#status translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-96 #label RES
#cross-references GB:M76595; NID:g182147; PIDN:AAA52393.1; PID:g553281
GENETICS
#gene GDB:EPOR
#map_position 19p13.3-19p13.2
#introns 39/1; 84/2; 143/1; 195/3; 247/1; 276/2; 305/3
CLASSIFICATION
#superfamily erythropoietin receptor; cytokine receptor
homology
KEYWORDS alternative splicing; cytokine receptor; glycoprotein;
transmembrane protein
FEATURE
1-24 #domain signal sequence #status predicted #label SIG\
25-508 #product erythropoietin receptor #status predicted
#label MAT\
52-250 #domain extracellular #status predicted #label EXT\
52-339 #domain cytokine receptor #status predicted #label CRS\
233-237 #region WSXS motif\
251-272 #domain transmembrane #status predicted #label TM\
273-508 #domain intracellular #status predicted #label INT\
#disulfide_bonds #status predicted\
52-62,91-107 #binding_site carbonylate (Asn) (covalent) #status
predicted\
SUMMARY #length 508 #molecular_weight 55065 #checksum 8723
Query Match 5.7%; Score 160; DB 1; Length 508;
Best Local Similarity 42.6%; Pred. NO. 7.36e-09;
Matches 26; Conservative 12; Mismatches 21; Indels 2; Gaps 2;
Db 248 LDPILTLISLVVILVLTVALSHRRALKQIKWIPSESEFEGLFTHKGNFOLW 307
QY 229 LSKFILLISLAILLMVSL-LLSLKLWRVKFLI-PSVPDPKSIFFGLFEHOGNFOEW 286
Db 308 L 308
QY 287 I 287
RESULT 10
ENTRY #type complete
TITLE interleukin-3 receptor beta chain precursor - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change
22-Jun-1999
ACCESSIONS A40091; A43022
REFERENCE A40091
#authors Itoh, N.; Yonehara, S.; Schreurs, J.; Gorman, D.M.; Maruyama,
K.; Ishii, A.; Yahara, I.; Arai, K.; Miyajima, A.
#journal Science (1990) 247:3243-327
#title Cloning of an interleukin-3 receptor gene: a member of a
distinct receptor gene family.
#cross-references MUID:90117145
#accession A40091
#status nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-878 #label ITO
REFERENCE A43022
#authors Gorman, D.M.
#submission submitted to GenBank, November 1989
#accession A43022
#molecule_type mRNA
#residues 1-815; 'Q', 817-878 #label GOR
#cross-references GB:M29855; NID:g198342; PIDN:AAA3295.1; PID:g309406
COMMENT In mice there are two classes of high-affinity IL-3 receptors. One
contains this IL-3-specific beta chain and the other contains the
beta chain also shared by high-affinity IL-5 and GM-CSF
receptors.
CLASSIFICATION #superfamily interleukin-3 receptor beta chain; cytokine
receptor; receptor homology
KEYWORDS cytokine receptor; duplication; transmembrane protein
FEATURE

```

```

1-22      #domain signal sequence #status predicted #label SIG\
23-878    #product interleukin-3 receptor beta chain #status
          predicted #label MAR\
23-440    #domain extracellular #status predicted #label EXT\
39-236    #domain cytokine receptor homology #label CRS1\
254-433   #domain cytokine receptor homology #label CRS2\
441-462   #domain transmembrane #status predicted #label TMM\
463-878   #domain intracellular #status predicted #label INT
SUMMARY   #length 878 #molecular-weight 97222 #checksum 2346

Query Match      4.1%; Score 115; DB 1; Length 878;
Best Local Similarity 32.4%; Pred. No. 2.80e-02;
Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 370 YIDHTFQYKKKSES-WKDSKTENLGRVNSMDLPDPTSYCARVRKPISD-YD-GI 426
| | :| | | :| | | | :| | :| | :| | :| | :| | :| | :| | :| |
QY 143 YGDLLEYQYRSPFDTEW-QSKQENTCNV-TIE-G-LDAEKCYSFWRVKAMEDVIGPDT 198
| | :| | | :| | | | :| | :| | :| | :| | :| | :| | :| | :| |
Db 427 W-SEWSNEYTW 436
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 199 YPSDMSSEVICW 209
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
ENTRY   #type complete
TITLE   cytokine receptor common beta chain precursor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE    30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change
02-Sep-1997
ACCESSIONS A39255
REFERENCE A39255
#authors Hayashida, K.; Kitamura, T.; Gorman, D.M.; Arai, K.; Yokota,
          T.; Miyajima, A.
#journal Proc. Natl. Acad. Sci. U.S.A. (1990) 87:9655-9659
#title Molecular cloning of a second subunit of the receptor for
          human granulocyte-macrophage colony-stimulating factor
          (GM-CSF): reconstitution of a high-affinity GM-CSF
          receptor.
#cross-references MUID:91088571
#accession A39255
#molecule_type mRNA
#residues 1-897 #label HAY
#cross-references GB:M38275
COMMENT The human high-affinity IL-3, IL-5, and GM-CSF receptors have
          ligand-specific alpha chains and share this common beta chain.
GENETICS
#gene GDB:CSF2RB
#cross-references GDB:126838; OMIM:138981
#map_position 22q13.1-22q13.1
CLASSIFICATION #superfamily interleukin-3 receptor beta chain; cytokine
          receptor homology
KEYWORDS alternative splicing; cytokine receptor; duplication;
          transmembrane protein
FEATURE
1-16      #domain signal sequence #status predicted #label SIG\
17-897    #product cytokine receptor common beta chain #status
          predicted #label MAR\
17-443    #domain extracellular #status predicted #label EXT\
35-232    #domain cytokine receptor homology #label CRS1\
250-431   #domain cytokine receptor homology #label CRS2\
444-460   #domain transmembrane #status predicted #label TMM\
461-897   #domain intracellular #status predicted #label INT
SUMMARY   #length 897 #molecular-weight 97323 #checksum 753

Query Match      4.1%; Score 113; DB 1; Length 897;
Best Local Similarity 31.9%; Pred. No. 5.15e-02;
Matches 22; Conservative 10; Mismatches 34; Indels 3; Gaps 3;

Db 372 FEIQRKDTAT-WKDSKTETLQNAHSMALPALEPSTRYWARVRVTSRTGYNGIWSNSE 430
| | :| | | :| | | | :| | :| | :| | :| | :| | :| | :| | :| |
QY 148 YEVOYRSPFDTEW-QSKQENTCNVTIEGLDA-EKCYSFWRVKAMEDVIGPDTYPSDSE 205
| | :| | | :| | | | :| | :| | :| | :| | :| | :| | :| | :| |
Db 431 ARSWDTESV 439
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

QY 206 VTCWQGEI 214
| :
| :
| :
RESULT 12
ENTRY   #type complete
TITLE   probable yop translocation J - Chlamydia trachomatis
          (serotype D, strain UW3/Cx)
ORGANISM #formal_name Chlamydia trachomatis
DATE    13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change
21-Nov-1998
ACCESSIONS H71497
REFERENCE A71570
#authors Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe,
          R.; Aravind, L.; Mitchell, W.P.; Olinger, L.; Tatusov,
          R.L.; Zhao, Q.; Koonin, E.V.; Davis, R.W.
#journal Science (1998) 282:754-759
#title Genome sequence of an obligate intracellular pathogen of
          humans: Chlamydia trachomatis.
#cross-references MUID:9900809
#accession H71497
#status preliminary
#molecule_type DNA
#residues 1-326 #label ARN
#cross-references GB:AE001327; GB:AE001273; NID:g3328999; PID:g3329000
#experimental_source serotype D, strain UW-3/Cx
GENETICS
#gene yscJ
SUMMARY #length 326 #molecular-weight 35553 #checksum 5831

Query Match      3.9%; Score 110; DB 2; Length 326;
Best Local Similarity 32.1%; Pred. No. 1.27e-01;
Matches 17; Conservative 14; Mismatches 20; Indels 2; Gaps 2;

Db 229 AKHSLTKRIVFVFLILLILSCGLLWVTKWHT-LISALGKTKGFDPPAPY 280
| | :| | | :| | | | :| | :| | :| | :| | :| | :| | :| | :| |
QY 225 PKPKLSKFLISSLAILLMVSLLLLSLWKLVRVKFLIPSPDPKSI-F-PGLF 276
| | :| | | :| | | | :| | :| | :| | :| | :| | :| | :| | :| |

RESULT 13
ENTRY   #type complete
TITLE   probable aspartate aminotransferase (tpaAT) - syphilis
          spirochete
ORGANISM #formal_name Treponema pallidum subsp. pallidum #common_name
          syphilis spirochete
DATE    24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change
17-Mar-1999
ACCESSIONS E71350
REFERENCE A71250
#authors Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.;
          Sutton, G.G.; Dodson, R.; Gwinn, M.; Hickey, E.K.; Clayton,
          R.; Ketchum, K.A.; Sodergren, E.; Hardham, J.M.; McLeod,
          M.P.; Salzberg, S.; Peterson, J.; Khalak, H.; Richardson,
          D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDonald,
          L.; Artiach, P.; Bowman, C.; Cotton, M.D.; Fujii, C.;
          Garland, S.; Hatch, B.; Horst, K.; Roberts, K.; Watthey,
          L.; Weidman, J.; Smith, H.O.; Venter, J.C.
#journal Science (1998) 281:375-388
#title Complete genome sequence of Treponema pallidum, the syphilis
          spirochete.
#cross-references MUID:98332770
#accession E71350
#status preliminary; nucleic acid sequence not shown;
          translation not shown
#molecule_type DNA
#residues 1-435 #label COL
#cross-references GB:AE001204; GB:AE000520; NID:g3322492; PID:g3322493
#experimental_source strain Nichols
GENETICS
#gene TP0223
SUMMARY #length 435 #molecular-weight 47304 #checksum 4156

Query Match      3.9%; Score 109; DB 2; Length 435;

```



```

QY 170 VTIEGLDAEKCYSEFWVRKAMEDVIGDTPSDWSEVTCWQGEIRDACAETPTPPKPL 229
Db 248 L-VGLSGAFGFIILVYLLINCRNTGPW-LKKVLKCNTPDPSKFFSOLSSSEHGDDVQKWL 305
QY 230 SKFILSSLAAILMVSLLLSLWK-LWRVKKFLIPSVDPKSIFFGLFEIHQGNFQEWIT 288

```

Search completed: Wed May 10 11:09:44 2000  
Job time : 28 secs.

\*\*\*\*\*  
M P S R C H  
\*\*\*\*\* (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 11:05:35 2000; MasPar time 192.75 Seconds  
133.451 Million cell updates/sec

Tabular output not generated.

Title: >US-09-376-430-2

Description: (1-371) from US09376430A.ppt (1 of 25)

Perfect Score: 2788

Sequence: 1 MGRLLVLLGCAAVFLGGWNA.....DVVTGGFTFVNDRSYVAL 371

Scoring table: PAM 150

Gap 11

Searched: 225878 seqs, 59334122 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: sptrembl12

1:sp\_archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
9:sp\_phage 10:sp\_plant 11:sp\_rodent 12:sp\_unclassified  
13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 47.769; Variance 88.048; scale 0.543

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	178	6.4	467	11	GFI-2	8.38e-13
2	172	4.7	383	11	IL-13 RECEPTOR ALPHA 2	1.82e-02
3	114	4.1	422	11	HYPOTHETICAL 45.4 KD P	6.88e-02
4	110	3.9	66	11	CITROCHROME P-450	1.30e-01
5	110	3.9	326	2	TOP TRANSLLOCATION J.	9.46e-01
6	108	3.9	427	4	INTERLEUKIN-13 RECEPTOR	9.46e-01
7	109	3.9	435	2	POTASSIUM ASPARTATE AMI	9.46e-01
8	109	3.9	435	2	ASPARTATE AMINOTRANSFER	9.46e-01
9	106	3.8	108	11	ALPHA 7C INTEGRIN (FRA	3.35e-01
10	105	3.8	359	11	C-C CHEMOKINE RECEPTOR	3.35e-01
11	105	3.8	634	6	SOMATOTROPIN RECEPTOR	2.45e-01
12	106	3.8	722	11	PREDICTED PROTEIN OF U	1.79e-01
13	107	3.8	787	11	INTERLEUKIN BETA-6 SUBUNI	2.45e-01
14	106	3.8	1106	11	INTERLEUKIN ALPHA 7 (ALPH	3.35e-01
15	105	3.8	1135	11	INTERLEUKIN SUBUNIT ALPH	2.45e-01
16	106	3.8	1167	4	INTERLEUKIN SUBUNIT ALPH	2.45e-01
17	106	3.8	1180	11	INTERLEUKIN ALPHA 7 PRECU	1.79e-01
18	107	3.8	1422	5	PFNS2L	8.45e-01
19	102	3.7	365	5	C23H5.1 PROTEIN.	8.45e-01
20	102	3.7	415	11	INTERLEUKIN-5 RECEPTOR	8.45e-01

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	467 AA.
ID	Q63216			
AC	Q63216			
DT	01-NOV-1996	(TREMBLrel. 01, Created)		
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)		
DT	01-NOV-1998	(TREMBLrel. 08, Last annotation update)		
DE	GFI-2.			
GN	GFI-2.			
OS	Rattus norvegicus (Rat).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
CC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
NC	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 95055995.			
RA	FLUBACHER M.M., BEAR S.E., TSICHLIS P.N.;			
RT	"Replacement of interleukin-2 (IL-2)-generated mitogenic signals by a			
RT	minik cell focus-forming (MCF) or xenotropic virus-induced IL-9-			
RT	dependent autocrine loop: implications for MCF virus-induced			
RT	leukemogenesis."			
RL	J. Virol. 68:7709-7716(1994).			
DR	EMBL; L36459; AAA63702.1; -			
SO	SEQUENCE 467 AA; 52017 MW; 0F3E477F CRC32;			

Query Match	6.4%	Score 178;	DB 11;	Length 467;
Best Local Similarity	30.0%	Pred. No. 8.38e-13;		
Matches	24;	Mismatches 29;	Indels 3;	Gaps 3;
	Conservative			
Db	272	ILVAVPIELLTGLIHF-LFRSLPKVKRIFVONVPSPEAFFHPLYSVHGDFOTWIGARR 330		
Qy	233	ILISSLAILLVSVLLSLWLKLV-RVKRELIPSVDPKSIFFGLFEIHQGNFQEWITDTQ 291		
Db	331	AGPQARQ-DGASAPSDSES 349		
Qy	292	NVAHLKMGAGQESGPEEP 311		
RESULT	2	PRELIMINARY;	PRT;	383 AA.
ID	O88786			
AC	O88786			
DT	01-NOV-1998	(TREMBLrel. 08, Created)		
DT	01-NOV-1998	(TREMBLrel. 08, Last sequence update)		
DT	01-NOV-1998	(TREMBLrel. 08, Last annotation update)		
DE	IL-13 RECEPTOR ALPHA 2.			
OS	Mus musculus (Mouse).			

[illegible]



```

ID O95646 PRELIMINARY; PRT; 427 AA.
AC O95646;
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMELrel. 12, Last annotation update)
DE INTERLEUKIN-13 RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA WADA M., HISANO T., KUWANO M.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81379; AAD00510.2; -.
KW Receptor.
SQ SEQUENCE 427 AA; 48737 MW; 52BF6F51 CRC32;

Query Match 3.9%; Score 108; DB 4; Length 427;
Best Local Similarity 22.5%; Pred. No. 1.30e-01;
Matches 64; Conservative 71; Mismatches 123; Indels 26; Gaps 26;

Db 122 EGPESAVTELQCIWNLSYMKCSWLPGRNTSPDNTYLYWHRSLKIHQCEN-IPREG 180
QY 23 QGGAEGV-QIILYFNLETQVITWASKY-SR-TNLTTHYRFGDEAYDQCTNYLLQEG 79
Db 181 QYFGCSFDLTQKVSFEQHSQIMVKONAGKIKPSFNIVP-LTSRVKDPDP-HIKNLSF 238
QY 80 HTSGC-L-LD-A-EQR-DD-ILYFSIR-N-GT-HPVFTASRWVYLYKPSPKHVR-FSW 129
Db 239 HNDLVQVENPQNFYSRCLFYEVYEVNNSQTHYFVYQEAACENPEPERNVENTSCFM 298
QY 130 HQDAVITCSD-LS-XGD-LLXEVQ-YRSPFDTEMOS-KQENTCNVTIEGLDAEKCSFW 184
Db 299 VPG-VLPDITNTVIRVK-TNKLCYEDDKLWNSWSEMSGKGRNSTLYITMLLIVPVI 356
QY 185 VRVKMEDVYGPDTYSDSEVTCWGEIRDA-CAETPPPRPKLSKFI-LISSLAILL 242
Db 357 AGAIIIVLLY-LKRLKIIIPPPDPGKIFKEMFGQNDTLHW 399
QY 243 MVSLLLSLWKLWRVKFLPSVDPDKSIFPGLFEIHQGNFQEW 286

RESULT 7 PRELIMINARY; PRT; 435 AA.
ID O06685;
AC O06685;
DT 01-JUL-1997 (TREMELrel. 04, Created)
DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)
DE PUTATIVE ASPARTATE AMINOTRANSFERASE TPAAT.
GN TPAAT.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
RN [1]
RP SEQUENCE FROM N.A.
RA SHEVCHENKO D.V., AKINS D.R., ROBINSON E., LI M., POPOVA T.G.;
RL MEDLINE; 97294457.
RT "Molecular characterization and cellular localization of TplRR, a processed leucine-rich repeat protein of Treponema pallidum, the syphilis spirochete."
RT J. Bacteriol. 179:3188-3195(1997).
DR EMBL; U73748; AAC45303.1; -.
KW Transferase; Aminotransferase.
SQ SEQUENCE 435 AA; 47285 MW; 7AF01306 CRC32;

Query Match 3.9%; Score 109; DB 2; Length 435;
Best Local Similarity 35.1%; Pred. No. 9.46e-02;
Matches 20; Conservative 11; Mismatches 21; Indels 5; Gaps 5;

Db 206 AVLVICDD-AYSGFEYEAASLMRGSFFARFAQAHNICALKIDGLTKEE-YA-WGLRV 259
QY 133 AVVTCSDLISGDLLEYVQY-RSPFDTEMOSKQENTCNVTIEGLDAEKCSFW-VRV 187

ID O63027 PRELIMINARY; PRT; 108 AA.
AC O63027;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)
DE -ALPHA 7C INTEGRIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RA SONG W.K., WANG W., SATO H., BIELSER D., KAUFMAN S.;
RL MEDLINE; 94171924.
RT "Expression of alpha 7 integrin cytoplasmic domains during skeletal muscle development: 7 alternate forms, conformational change, and homologues with serine/threonine kinases and tyrosine phosphatases."
RL J. Cell Sci. 106:1139-1152(1993).
DR EMBL; X74294; CAA52347.1;
DR PFAM; PF00357; integrin_A; 1.
KW Integrin.
SQ SEQUENCE 108 AA; 12016 MW; 1A077241 CRC32;

```

```

RESULT 8 PRELIMINARY; PRT; 435 AA.
ID O83252;
AC O83252;
DT 01-NOV-1998 (TREMELrel. 08, Created)
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)
DE ASPARTATE AMINOTRANSFERASE (TPAAT).
GN TP0223.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
RN [1]
RP SEQUENCE FROM N.A.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.;
RX MEDLINE; 98332770.
RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.;
RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.;
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.;
RA McDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.;
RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.;
RA VENTER J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis spirochete."
RL Science 281:375-388(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.;
RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.;
RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.;
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.;
RA McDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.;
RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.;
RA VENTER J.C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE001204; AAC65212.1; -.
KW TIGR; TP0223; -.
KW Transferase; Aminotransferase.
SQ SEQUENCE 435 AA; 47304 MW; 0B31A6E1 CRC32;

Query Match 3.9%; Score 109; DB 2; Length 435;
Best Local Similarity 35.1%; Pred. No. 9.46e-02;
Matches 20; Conservative 11; Mismatches 21; Indels 5; Gaps 5;

Db 206 AVLVICDD-AYSGFEYEAASLMRGSFFARFAQAHNICALKIDGLTKEE-YA-WGLRV 259
QY 133 AVVTCSDLISGDLLEYVQY-RSPFDTEMOSKQENTCNVTIEGLDAEKCSFW-VRV 187

RESULT 9 PRELIMINARY; PRT; 108 AA.
ID O63027;
AC O63027;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)
DE -ALPHA 7C INTEGRIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RA SONG W.K., WANG W., SATO H., BIELSER D., KAUFMAN S.;
RL MEDLINE; 94171924.
RT "Expression of alpha 7 integrin cytoplasmic domains during skeletal muscle development: 7 alternate forms, conformational change, and homologues with serine/threonine kinases and tyrosine phosphatases."
RL J. Cell Sci. 106:1139-1152(1993).
DR EMBL; X74294; CAA52347.1;
DR PFAM; PF00357; integrin_A; 1.
KW Integrin.
SQ SEQUENCE 108 AA; 12016 MW; 1A077241 CRC32;

```

---

```

Query Match          3.88; Score 106; DB 11; Length 1106;
Best Local Similarity 59.1%; Pred. NO. 2.45e-01;
Matches 13; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

```

Search completed: Wed May 10 11:08:58 2000  
Job time : 203 secs.

0.



Thu May 11 06:49:26 2000

\*\*\*\*\*  
MPELRH  
\*\*\*\*\*  
(TM)  
\*\*\*\*\*  
Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd  
MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:18:00 2000; MasPar time 13.27 Seconds  
Tabular output not generated. 622.771 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (23-371) from US09376430A.pep (2 of 25)  
Perfect Score: 2401  
Sequence: 1 QGGAEGVOIILFNLETV.....DVTIGGTFVMDRSYVAL 349  
Scoring table: PAM 150  
Gap 11  
Searched: 189963 seqs, 23686106 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: a-geneseq35  
1:geneseq  
Statistics: Mean 34.908; Variance 149.576; scale 0.233

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	200	7.7	482	1 W31646	Human cytokine recepto	1.49e-07
2	195	7.5	230	1 R82934	Interleukin 4 componen	4.01e-07
3	195	7.5	230	1 R47151	IL-2 receptor gamma ch	4.01e-07
4	195	7.5	252	1 R47150	IL-2 receptor gamma ch	4.01e-07
5	195	7.5	347	1 R47149	IL-2 receptor gamma ch	4.01e-07
6	195	7.5	369	1 R47148	IL-2 receptor gamma ch	4.01e-07
7	188	7.2	369	1 R59094	Murine IL-2K gamma.	1.59e-06
8	182	7.0	500	1 W64057	Human IL-9 receptor pr	5.15e-06
9	176	6.8	501	1 W64055	Human IL-9 receptor pr	1.85e-05
10	175	6.7	501	1 W64056	Human IL-9 receptor pr	2.01e-05
11	163	6.3	507	1 R47517	MEL EPO receptor.	2.02e-04
12	163	6.3	507	1 R69502	Mouse erythropoietin r	2.02e-04
13	163	6.3	507	1 R05111	EPO receptor sequence	3.57e-04
14	160	6.2	508	1 R69503	Human erythropoietin r	3.57e-04
15	160	6.2	508	1 R05112	EPO receptor.	3.57e-04
16	160	6.2	508	1 R47518	Human erythropoietin r	3.57e-04
17	150	6.2	508	1 W35294	Murine IL-13 binding c	9.37e-02
18	130	5.0	383	1 R09821	Human IL-13 binding c	1.13e+00
19	116	4.5	426	1 R06846	Murine interleukin-12 r	1.60e+00
20	114	4.4	551	1 R06846	Human interleukin-12 r	1.34e+00
21	114	4.4	551	1 R78613	Expression vector pME1	1.34e+00
22	115	4.4	576	1 R78613	Expression vector pME1	1.34e+00
23	115	4.4	596	1 R78616	Expression vector pME1	1.34e+00

24	115	4.4	600	1 R78610	Expression vector pME1	1.34e+00
25	115	4.4	600	1 R78610	Fas antigen #1.	1.34e+00
26	115	4.4	878	1 R78608	Murine IL-3 receptor b	1.34e+00
27	111	4.3	237	1 R22229	Truncated human growth	2.68e+00
28	111	4.3	269	1 R20545	Soluble human growth h	2.68e+00
29	112	4.3	551	1 R07506	IL-2R beta chain.	2.26e+00
30	112	4.3	592	1 R25227	Fas antigen #2.	2.26e+00
31	113	4.3	897	1 R20982	Sequence of beta-chain	1.90e+00
32	110	4.2	237	1 R24274	Truncated human growth	3.18e+00
33	110	4.2	551	1 W39210	Human interleukin-2 re	3.18e+00
34	110	4.2	637	1 P92108	Human growth hormone r	3.18e+00
35	110	4.2	638	1 R33394	Human growth hormone r	3.18e+00
36	109	4.2	878	1 R32529	Fas sequence from AIC2	3.78e+00
37	107	4.1	426	1 W09822	Human interleukin-12 r	5.31e+00
38	107	4.1	427	1 W24973	Human interleukin-13 a	5.31e+00
39	107	4.1	459	1 R08329	Human IL-7 receptor cl	5.31e+00
40	106	4.1	1073	1 W37371	Human ST receptor prot	6.29e+00
41	106	4.1	1073	1 W32063	Human ST receptor prot	6.29e+00
42	105	4.0	572	1 Y04954	Mycobacterium species	7.45e+00
43	103	4.0	638	1 P81326	Human growth hormone r	1.04e+01
44	104	4.0	1193	1 Y07728	Armenian hamster alpha	8.81e+00
45	100	3.8	919	1 R75504	Human mammary carcino	1.72e+01

ALIGNMENTS

RESULT 1  
ID W31646 standard; Protein: 482 AA.

AC	W31646;				
DE	21-MAY-1998 (first entry)				
DE	Human cytokine receptor gc chain-Ig fusion protein.				
DE	Cytokine receptor; gamma common chain; gc chain; human;				
KW	blocking agent; monoclonal antibody; Cp.B8; immunological disease;				
KW	myasthenia gravis; rheumatoid arthritis; lupus; multiple sclerosis;				
KW	insulin-dependent diabetes; inflammatory bowel disease;				
KW	sympathetic ophthalmia; uveitis; allergy; asthma; infection;				
KW	graft versus host disease; psoriasis; immunosuppressive; therapy.				
OS	Chimeric - Homo sapiens.				
PH	Key				
FT	Protein				
FT	1..254				
FT	/note= "gc chain N-terminal region"				
FT	255..482				
FT	/note= "IgG1 constant region"				
FT	255..264				
FT	/note= "IgG1 hinge region"				
FT	264..482				
FT	/note= "IgG1 CH2 and CH3 constant domains1"				
FT	Domain				
FT	W09743416-A1.				
PN	20-NOV-1997.				
PD	09-MAY-1997; U07870.				
PP	10-MAY-1996; US-017466.				
PP	(BIOJ ) BIOGEN INC.				
PA	Benjamin CD, Burkly LC, Hession C, Whitty A;				
PI	WPI; 98-008885/01.				
DR	N-PSDB; T97439.				
DR	Blocking agents of the gamma common chain of cytokine receptors - for				
PT	particularly monoclonal antibodies, used to induce T cell anergy for				
PT	treatment of immunological diseases				
PT	Example 1: Page 79-80; 111pp; English.				
PS	This polypeptide comprises a fusion between the N-terminal 254				
PS	amino acids of the human mature cytokine receptor gamma common (gc)				
CC	chain and the hinge region and CH2 and CH3 constant domains of				
CC	human IgG1. The fusion was expressed from clone pLB001 (see				
CC	T97439) in COS-7 cells, and used to generate murine anti-human gc				
CC	specific monoclonal antibodies (MAbs), including cp.B8 produced by				
CC	hybridoma MTCC RB 12107. The invention provides compositions and				
CC	methods for inhibiting cytokine signalling using gc chain blocking				
CC	agents for the treatment of immunological diseases such as				
CC	myasthenia gravis, rheumatoid arthritis, lupus, multiple sclerosis,				
CC	insulin-dependent diabetes, inflammatory bowel disease, sympathetic				
CC	ophthalmia, uveitis, allergy, asthma, parasitic infection, graft				
CC	vs. host disease or psoriasis. A preferred gc blocking agent is				
CC	MAB CP.B8 or its Fab fragment (see also W31647-48).				

SQ Sequence 482 AA;

Query Match 7.7%; Score 200; DB 1; Length 482;  
 Best Local Similarity 25.7%; Pred. No. 1.49e-07;  
 Matches 53; Conservative 53; Mismatches 88; Indels 12; Gaps 12;

Db 59 EVQCFVFNVEYMNCTWNSSEPPQTNLTLYHYKNSDNDKVKQKSHYLFSEITSGCOLQ 118  
 QY 31 QIQIYFNLETVQVTWNSKYSR-TNLTFFHYRF-NGD-EAYDOCTNYLLQEGHTSGCLLD 87  
 Db 119 KKEIHLQYTFVVOLODPPRPRQATQMLKQNLVWPAPENLTLHKLSQLELNNNRF 178  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYLL-KPSSPKHVRP-SWQDAVTVTCSDLS 142  
 Db 179 LNHCLHLYQYRTDWSHTEQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 237  
 QY 143 YGD-LLYEVQVRSFDFTEW-QSKQENTCNVTIEGLDAEKCYSFVVRVKAMEDVYGPDTYP 200  
 Db 238 SEWSHPHIGSNTSKENVDKTHTCPP 263  
 QY 201 SDWSEVTCWQGEIRDAETPT-PP 225

## RESULT 2

ID R82934 standard; Protein; 230 AA.

DT 26-FEB-1996 (first entry)  
 DE Interleukin 4 component common to the IL-2 receptor gamma chain.  
 KW Interleukin-4; IL-4; gamma chain component; immunosuppressants;  
 KW anti-allergy agent; signal transmission inhibitor; autoimmune;  
 KW disease; anti-inflammatory; anaphylactic shock; bronchial asthma;  
 KW Interleukin-2; IL-2; atopic dermatitis; urticaria.  
 OS Homo sapiens.  
 PN J07149662-A.  
 PD 13-JUN-1995.

PF 07-SEP-1994; 213706.

PR 08-SEP-1993; JP-223574.

PA (AJIN) AJINOMOTO KK.

PA (SUGA) SUGAMURA K.

DR WPI; 95-243601/32.

DR N-PSDB; T04952.

PT Novel interleukin-4 receptor monoclonal antibodies inhibit signal transmission - useful as immunosuppressants and anti-allergy agents.  
 PS Example 1; Page 9; lipp; Japanese.

CC T04952 encodes R82934 a component of the IL-4 receptor common to the IL-2 receptor gamma chain molecule, which was used to generate anti-IL-4 receptor monoclonal antibodies (mAbs). The mAbs (IL-4 signal transmission inhibitors) can be used as immunosuppressants and anti-allergy agents, for the treatment of autoimmune and chronic inflammatory diseases, e.g. anaphylactic shock, bronchial asthma, atopic dermatitis and urticaria.

SQ Sequence 230 AA;

Query Match 7.5%; Score 195; DB 1; Length 230;  
 Best Local Similarity 25.9%; Pred. No. 4.01e-07;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;

Db 38 EVQCFVFNVEYMNCTWNSSEPPQTNLTLYHYKNSDNDKVKQKSHYLFSEITSGCOLQ 97  
 QY 31 QIQIYFNLETVQVTWNSKYSR-TNLTFFHYRF-NGD-EAYDOCTNYLLQEGHTSGCLLD 87  
 Db 98 KKEIHLQYTFVVOLODPPRPRQATQMLKQNLVWPAPENLTLHKLSQLELNNNRF 157  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYLL-KPSSPKHVRP-SWQDAVTVTCSDLS 142  
 Db 158 LNHCLHLYQYRTDWSHTEQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 216  
 QY 143 YGD-LLYEVQVRSFDFTEW-QSKQENTCNVTIEGLDAEKCYSFVVRVKAMEDVYGPDTYP 200  
 Db 217 SEWSHPHWH 225  
 QY 201 SDWSEVTCW 209

## RESULT 3

ID R47151 standard; Protein; 230 AA.

AC R47151;  
 DT 13-JUN-1994 (first entry)  
 DE IL-2 receptor gamma chain.  
 KW Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;  
 KW rheumatoid arthritis; transplant rejection; primer;  
 KW polymerase chain reaction; PCR; amplification.  
 OS Homo sapiens.  
 PN EP-578932-A.  
 PD 19-JAN-1994.

PF 22-APR-1993; 106561.

PR 23-APR-1992; JP-104947.

PA (AJIN) AJINOMOTO KK.

PA (SUGA) SUGAMURA K.

PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;

PI Suzuki M, Takeshita T;

DR WPI; 94-017546/03.

P-PSDB; Q54831.

PT DNA and protein sequences of IL-2 gamma chain - useful as immune regulatory agents for treatment of e.g. rheumatoid arthritis and transplant rejection

PS Disclosure; Page 22-23, 35-36; 50pp; English.

CC The human IL-2 receptor gamma chain preform (R47148), including the signal peptide, is encoded by the sequence given in Q54828. The mature protein (R47149) is encoded by sequence Q54829. A soluble form of IL-2 receptor gamma chain (R47150) is encoded by Q54830, CC while a soluble form suitable for expression in prokaryotes (R47151) is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-terminal sequence of IL-2 receptor gamma chain, and are used to CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27 CC are used to obtain the protein given in R47151.

SQ Sequence 230 AA;

Query Match 7.5%; Score 195; DB 1; Length 230;

Best Local Similarity 25.9%; Pred. No. 4.01e-07;

Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;

Db 37 EVQCFVFNVEYMNCTWNSSEPPQTNLTLYHYKNSDNDKVKQKSHYLFSEITSGCOLQ 96  
 QY 31 QIQIYFNLETVQVTWNSKYSR-TNLTFFHYRF-NGD-EAYDOCTNYLLQEGHTSGCLLD 87  
 Db 97 KKEIHLQYTFVVOLODPPRPRQATQMLKQNLVWPAPENLTLHKLSQLELNNNRF 156  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYLL-KPSSPKHVRP-SWQDAVTVTCSDLS 142  
 Db 157 LNHCLHLYQYRTDWSHTEQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 215  
 QY 143 YGD-LLYEVQVRSFDFTEW-QSKQENTCNVTIEGLDAEKCYSFVVRVKAMEDVYGPDTYP 200  
 Db 216 SEWSHPHWH 224  
 QY 201 SDWSEVTCW 209

## RESULT 4

ID R47150 standard; Protein; 252 AA.

AC R47150;  
 DT 13-JUN-1994 (first entry)  
 DE IL-2 receptor gamma chain.  
 KW Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;  
 KW rheumatoid arthritis; transplant rejection; primer;  
 KW polymerase chain reaction; PCR; amplification; ss.  
 OS Homo sapiens.

PN Key Location/Qualifiers

FT peptide 1..22

FT /label= Sig\_peptide

PN EP-578932-A.

PD 19-JAN-1994.

PF 22-APR-1993; 106561.

PR 23-APR-1992; JP-104947.

PA (AJIN) AJINOMOTO KK.

Thu May 11 06:49:26 2000

(SUGA/) SUGAMURA K.  
 PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
 PI Suzuki M, Takeshita T;  
 DR WPI; 94-017546/03.  
 DR P-PSDB; Q54830.  
 DR DNA and protein sequences of IL-2 gamma chain - useful as immune  
 PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
 PT transplant rejection  
 PS Disclosure; Page 21-22, 34-35; 50pp; English.  
 CC The human IL-2 receptor gamma chain preform (R47148), including the  
 CC signal peptide, is encoded by the sequence given in Q54828. The  
 CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
 CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
 CC while a soluble form suitable for expression in prokaryotes (R47151)  
 CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
 CC terminal sequence of IL-2 receptor gamma chain, and are used to  
 CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
 CC are used to obtain the protein given in R47151.  
 SQ Sequence 252 AA;  
 Query Match 7.5%; Score 195; DB 1; Length 252;  
 Best Local Similarity 25.9%; Pred. No. 4.01e-07;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 59 EVQCFVFNVEYMNCTWNSSEPOPTNLTLYHYKNSDNDKVKCSHYLSEITSGCOLQ 118  
 QY 31 QIQIIFYNLETQVQVWNASKYSR-TNLTFHYRF-NGD-EAYDOCTNYLLOEGHTSGCLLD 87  
 Db 119 KKEIHLVQYTFVVOLODPREPRQATOMLKLQNLVWPAPENLTLLKLSSEQLNNRNF 178  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRWM-VYIL-KPSSPKHVRF-SWQDAVTVTCSDLS 142  
 Db 179 LNHCLHLVQYTRDWDHSHWTEQSDYRHKFSLPSVDGQKRYTFVRVSR-FNPLCGSAQHW 237  
 QY 143 YGD-LLEYEQVRSFPDTEW-QSKQENTCNVTIEGLDAEKYSFVWVRKAMEDVYGPDTYP 200  
 Db 238 SEWSHPHWH 246  
 QY 201 SDWSEVTCW 209  
 RESULT 5  
 ID R47149 standard; Protein; 347 AA.  
 AC R47149;  
 DE 13-JUN-1994 (first entry)  
 KW Interleukin-2 receptor gamma chain  
 KW rheumatoid arthritis; transplant rejection; primer;  
 KW polymerase chain reaction; PCR; amplification.  
 OS Homo sapiens.  
 PN EP-578932-A.  
 PD 19-JAN-1994. 106561.  
 PF 22-APR-1993; JP-104947.  
 PR 23-APR-1992; JP-104947.  
 PA (AJIN ) AJINOMOTO KK.  
 PA (SUGA/) SUGAMURA K.  
 PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
 PI Suzuki M, Takeshita T;  
 DR WPI; 94-017546/03.  
 DR P-PSDB; Q54828.  
 DR DNA and protein sequences of IL-2 gamma chain - useful as immune  
 PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
 PT transplant rejection  
 PS Disclosure; Page 16-17, 29-30; 50pp; English.  
 CC The human IL-2 receptor gamma chain preform (R47148), including the  
 CC signal peptide, is encoded by the sequence given in Q54828. The  
 CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
 CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
 CC while a soluble form suitable for expression in prokaryotes (R47151)  
 CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
 CC terminal sequence of IL-2 receptor gamma chain, and are used to  
 CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
 CC are used to obtain the protein given in R47151.  
 SQ Sequence 347 AA;  
 Query Match 7.5%; Score 195; DB 1; Length 347;  
 Best Local Similarity 25.9%; Pred. No. 4.01e-07;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 59 EVQCFVFNVEYMNCTWNSSEPOPTNLTLYHYKNSDNDKVKCSHYLSEITSGCOLQ 118  
 QY 31 QIQIIFYNLETQVQVWNASKYSR-TNLTFHYRF-NGD-EAYDOCTNYLLOEGHTSGCLLD 87  
 Db 119 KKEIHLVQYTFVVOLODPREPRQATOMLKLQNLVWPAPENLTLLKLSSEQLNNRNF 178  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRWM-VYIL-KPSSPKHVRF-SWQDAVTVTCSDLS 142  
 Db 179 LNHCLHLVQYTRDWDHSHWTEQSDYRHKFSLPSVDGQKRYTFVRVSR-FNPLCGSAQHW 237  
 QY 143 YGD-LLEYEQVRSFPDTEW-QSKQENTCNVTIEGLDAEKYSFVWVRKAMEDVYGPDTYP 200  
 Db 238 SEWSHPHWH 246  
 QY 201 SDWSEVTCW 209  
 RESULT 5  
 ID R47149 standard; Protein; 347 AA.  
 AC R47149;  
 DE 13-JUN-1994 (first entry)  
 KW Interleukin-2 receptor gamma chain  
 KW rheumatoid arthritis; transplant rejection; primer;  
 KW polymerase chain reaction; PCR; amplification.  
 OS Homo sapiens.  
 PN EP-578932-A.  
 PD 19-JAN-1994. 106561.  
 PF 22-APR-1993; JP-104947.  
 PR 23-APR-1992; JP-104947.  
 PA (AJIN ) AJINOMOTO KK.  
 PA (SUGA/) SUGAMURA K.  
 PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
 PI Suzuki M, Takeshita T;  
 DR WPI; 94-017546/03.  
 DR P-PSDB; Q54828.  
 DR DNA and protein sequences of IL-2 gamma chain - useful as immune  
 PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
 PT transplant rejection  
 PS Disclosure; Page 16-17, 29-30; 50pp; English.  
 CC The human IL-2 receptor gamma chain preform (R47148), including the  
 CC signal peptide, is encoded by the sequence given in Q54828. The  
 CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
 CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
 CC while a soluble form suitable for expression in prokaryotes (R47151)  
 CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
 CC terminal sequence of IL-2 receptor gamma chain, and are used to  
 CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
 CC are used to obtain the protein given in R47151.  
 SQ Sequence 347 AA;  
 Query Match 7.5%; Score 195; DB 1; Length 347;  
 Best Local Similarity 25.9%; Pred. No. 4.01e-07;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;

SQ Sequence 347 AA;  
 Query Match 7.5%; Score 195; DB 1; Length 347;  
 Best Local Similarity 25.9%; Pred. No. 4.01e-07;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 37 EVQCFVFNVEYMNCTWNSSEPOPTNLTLYHYKNSDNDKVKCSHYLSEITSGCOLQ 96  
 QY 31 QIQIIFYNLETQVQVWNASKYSR-TNLTFHYRF-NGD-EAYDOCTNYLLOEGHTSGCLLD 87  
 Db 97 KKEIHLVQYTFVVOLODPREPRQATOMLKLQNLVWPAPENLTLLKLSSEQLNNRNF 156  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRWM-VYIL-KPSSPKHVRF-SWQDAVTVTCSDLS 142  
 Db 157 LNHCLHLVQYTRDWDHSHWTEQSDYRHKFSLPSVDGQKRYTFVRVSR-FNPLCGSAQHW 215  
 QY 143 YGD-LLEYEQVRSFPDTEW-QSKQENTCNVTIEGLDAEKYSFVWVRKAMEDVYGPDTYP 200  
 Db 216 SEWSHPHWH 224  
 QY 201 SDWSEVTCW 209  
 RESULT 6  
 ID R47148 standard; Protein; 369 AA.  
 AC R47148;  
 DE 13-JUN-1994 (first entry)  
 KW Interleukin-2 receptor gamma chain  
 KW rheumatoid arthritis; transplant rejection; primer; PCR;  
 KW polymerase chain reaction; amplification; ss.  
 OS Homo sapiens.  
 PN EP-578932-A.  
 PD 19-JAN-1994. 106561.  
 PF 22-APR-1993; JP-104947.  
 PR 23-APR-1992; JP-104947.  
 PA (AJIN ) AJINOMOTO KK.  
 PA (SUGA/) SUGAMURA K.  
 PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
 PI Suzuki M, Takeshita T;  
 DR WPI; 94-017546/03.  
 DR P-PSDB; Q54828.  
 DR DNA and protein sequences of IL-2 gamma chain - useful as immune  
 PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
 PT transplant rejection  
 PS Disclosure; Page 16-17, 29-30; 50pp; English.  
 CC The human IL-2 receptor gamma chain preform (R47148), including the  
 CC signal peptide, is encoded by the sequence given in Q54828. The  
 CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
 CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
 CC while a soluble form suitable for expression in prokaryotes (R47151)  
 CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
 CC terminal sequence of IL-2 receptor gamma chain, and are used to  
 CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
 CC are used to obtain the protein given in R47151.  
 SQ Sequence 369 AA;  
 Query Match 7.5%; Score 195; DB 1; Length 369;  
 Best Local Similarity 25.9%; Pred. No. 4.01e-07;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 59 EVQCFVFNVEYMNCTWNSSEPOPTNLTLYHYKNSDNDKVKCSHYLSEITSGCOLQ 118  
 QY 31 QIQIIFYNLETQVQVWNASKYSR-TNLTFHYRF-NGD-EAYDOCTNYLLOEGHTSGCLLD 87  
 Db 119 KKEIHLVQYTFVVOLODPREPRQATOMLKLQNLVWPAPENLTLLKLSSEQLNNRNF 178  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRWM-VYIL-KPSSPKHVRF-SWQDAVTVTCSDLS 142  
 Db 179 LNHCLHLVQYTRDWDHSHWTEQSDYRHKFSLPSVDGQKRYTFVRVSR-FNPLCGSAQHW 237

143 YGD-LLYEYVQYRSPFDTEW-QSKOENTCNVTIEGLDAEKYCFWVRVKAMEDVYGPDTYP 200

238 SEWSHPHWH 246

201 SDNSEVTCW 209

RESULT 7  
ID R59094 standard; Protein; 369 AA.  
AC R59094;  
DE Murine IL-2R gamma  
KW Murine IL2-R gamma; X-linked severe combined immunodeficiency;  
OS XSCID; interleukin.  
FH Mus musculus.  
FT Key  
FT peptide  
FT domain  
FT /note= "signal peptide"  
FT /note= "transmembrane domain"  
FT misc\_difference 331  
FT /note= "Corresponding codon CAG"  
FT modified\_site 71..73  
FT /label= N-glycosylation\_site  
FT modified\_site 75..77  
FT /label= N-glycosylation\_site  
FT modified\_site 84..86  
FT /label= N-glycosylation\_site  
FT modified\_site 96..98  
FT /label= N-glycosylation\_site  
FT modified\_site 159..161  
FT /label= N-glycosylation\_site  
FT modified\_site 255..257  
FT /label= N-glycosylation\_site  
PN W09420641-A.  
PD 15-SEP-1994.  
PF 10-MAR-1994; U02891.  
PR 12-MAR-1993; US-031143.  
PR 14-SEP-1993; US-121435.  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PI Leonard WJ, McBride WO, Noguchi M;  
DR WPI; 94-303046/37.  
DR N-PSDB; Q71977.

Diagnosis of X-linked severe combined immunodeficiency (XSCID) -  
comprises detecting mutated IL-2R gamma gene, also vectors and  
transgenic animals containing the mutated gene  
Example 1: Fig 7; 98pp; English.  
Q71977 is the DNA sequence of murine IL-2R gamma R59094.  
This was used in the development of a claimed method for the  
diagnosis of X-linked severe combined immunodeficiency (XSCID),  
in female carriers and male sufferers.  
SQ Sequence 369 AA;

Query Match 7.2%; Score 188; DB 1; Length 369;  
Best Local Similarity 26.8%; Pred. No. 1.59e-06;  
Matches 51; Conservative 52; Mismatches 75; Indels 12; Gaps 11;

Db 59 EVQCFVNIETWNCWSSPEQATNLHLRYKVDNNTQECSHYLFSEKTSQCQIQ 118  
QY 31 QIQIYFNLETVQVNTWASKYSR-TNLTFFHYRFN-GDE-AYDOCTNYLLQEGTSGCLLD 87  
Db 119 KEDIQYQTFVQIQDQPKORRAVQKLNQNLVTPRAPENLTLSNLSQLELWKRSH 178  
QY 88 AEQRDILLYFSIR-NGRH-PVFTASRW-VYIL-KPSSPKHVRFS-WHQDAVTV--TCS 140  
Db 179 IKERCLQYLYQYRNRDRSWTELINHEPRFSLPSVDLKRITFRVGRY-NPICGSSQ 237  
QY 141 LSYGDLLEYQYRSPFDTEW-QSKOENTCNVTIEGLDAEKYCFWVRVKAMEDVYGPDTY 199  
Db 238 WSKWSQVPHW 247  
QY 200 PSDNSEVTCW 209

RESULT 8  
ID W64057 standard; Protein; 500 AA.  
AC W64057;  
DT 02-OCT-1998 (first entry)  
DE Human IL-9 receptor protein variant #2.  
KW Interleukin-9 receptor; IL-9R; Asthma associated factor 2; AAF2;  
KW soluble; bronchial hyper-responsiveness; detection; diagnosis; human;  
KW eosinophil; HL60 cell; antagonist.  
OS Homo sapiens.  
PN W09824904-A2.  
PD 11-JUN-1998.  
PF 02-DEC-1997; U21992.  
PR 01-DEC-1997; US-032224.  
PR 02-DEC-1996; US-032224.  
PA (MAGA-) MAGAININ PHARM INC.  
PI Grasso L, Holroyd KJ, Levitt RC, Nicolaides NC;  
DR WPI; 98-348150/30.  
PT Human interleukin-9 receptor variant(s) useful in treating asthma -  
and related disorders e.g. bronchial hyper-responsiveness and for  
diagnosing greater or alternatively less susceptibility to these  
conditions  
PS Disclosure; Page 8; 81pp; English.  
CC This sequence represents a human interleukin-9 (IL-9) receptor variant  
CC (also known as Asthma Associated Factor 2 or AAF2) which is used in the  
CC construction of protein variants which are soluble and can be  
CC administered to humans to alleviate asthma and related disorders e.g.  
CC bronchial hyper-responsiveness. The DNA molecules are useful to detect  
CC or diagnose susceptibility to such conditions. Cells used in this method  
CC may be e.g. eosinophils and HL60 cells. Cells lacking human IL-9 receptor  
CC can be transfected with the DNA and used to identify IL-9 pathway  
CC antagonists. The nucleic acids can also be used to design probes to  
CC detect other IL-9 receptor variants.  
SQ Sequence 500 AA;

Query Match 7.0%; Score 182; DB 1; Length 500;  
Best Local Similarity 26.9%; Pred. No. 5.15e-06;  
Matches 61; Conservative 55; Mismatches 87; Indels 24; Gaps 20;

Db 157 LLSYELAFKKQ-EENAEQHRDHIVGVTWLLEAFELDPGFIHEARLRVOMATLEDDYVE 215  
QY 146 LL-YEYVQYRSPFDTEWQSKO-ENTCNVT--I-EGLDAEKYCF--WVRV-KA-MED-VYG 195  
Db 216 EERYTGQWSE--WSQPVCFQAPQROGFLPPKWPQCN-TIVAVSIPLLLTGPTYL-LFK 270  
QY 196 PDYPSDSEVTCWQGE-IRDAETP-TTPKPKLSKFLISSLAILMLVSLLLSLWK 253  
Db 271 LSPRVKRIYQNVPSAMPFFQPLYSVHNGNFQTMGAHRAGVLLSQDCAGTPQGALEPCV 330  
QY 254 LW-RVKKFLIPSPVDPKSIFFGLFEIHQGNFQEWITDQNVHL-HKMGAEQES-GPEE 310  
Db 331 QEATALLTCGPAPRWKSVALEEEQEGPG-T-RLPGN-LSSSEDLVPAG 374  
QY 311 PLVVQLAKTEASPRMLDPQTEKEASGSLQLPHOPLOGGVVVTIG 357

RESULT 9  
ID W64055 standard; Protein; 501 AA.  
AC W64055;  
DT 02-OCT-1998 (first entry)  
DE Human IL-9 receptor protein.  
KW Interleukin-9 receptor; IL-9R; Asthma associated factor 2; AAF2;  
KW soluble; bronchial hyper-responsiveness; detection; diagnosis; human;  
KW eosinophil; HL60 cell; antagonist.  
OS Homo sapiens.  
PN W09824904-A2.  
PD 11-JUN-1998.  
PF 02-DEC-1997; U21992.  
PR 01-DEC-1997; US-032224.  
PR 02-DEC-1996; US-032224.  
PA (MAGA-) MAGAININ PHARM INC.  
PI Grasso L, Holroyd KJ, Levitt RC, Nicolaides NC;



Thu May 11 06:49:26 2000

WPI; 98-348150/30.  
N-PSDB; V44088.  
Human interleukin-9 receptor variant (s) useful in treating asthma -  
and related disorders e.g. bronchial hyper-responsiveness and for  
diagnosing greater or alternatively less susceptibility to these  
conditions  
Disclosure; Fig 2; 81pp; English.  
This sequence represents a human interleukin-9 (IL-9) receptor (also  
known as Asthma Associated Factor 2 or AAF2) which is used in the  
construction of protein variants which are soluble and can be  
administered to humans to alleviate asthma and related disorders e.g.  
bronchial hyper-responsiveness. The DNA molecules are useful to detect  
or diagnose susceptibility to such conditions. Cells used in this method  
may be e.g. eosinophils and HL60 cells. Cells lacking human IL-9 receptor  
can be transfected with the DNA and used to identify IL-9 pathway  
antagonists. The nucleic acids can also be used to design probes to  
detect other IL-9 receptor variants.  
Sequence 501 AA;  
Query Match 6.8%; Score 176; DB 1; Length 501;  
Best Local Similarity 27.6%; Pred. No. 1.65e-05;  
Matches 50; Conservative 43; Mismatches 74; Indels 14; Gaps 12;  
Db 203 LRQMATELDDVBEERYTGWSE---WSQPVCFQAPQROGLIPPGWPGN-TLVAISI 258  
QY 183 FWRVKAMED-VYGPDTVPDSEVTCWQGE-IRDAETP-TPPKPKLSKFLISLSLA 239  
Db 259 FLLLTGPTYL-LFKLSPRVKRIFQNVPSAMFFQPLYSVHNGFQWMAHRAAGVLLSQ 317  
QY 240 ILLMVSLLLSLWKLW-RVKKFLIPSPDPKSIFFGLFEIHQGNFQEWITDTQNVHL-H 297  
Db 318 DCAGTPQGALEPCVQEAATLLTCGPAPKWSVALEEEQEGPG-T-RLPGN-LSSEDLVPA 374  
QY 298 KMGAEQES-GPEELVVLQAKTEASPRMLDPQTEKEASGSLQLPHQLOGGDVVTI 356  
Db 375 G 375  
QY 357 G 357  
RESULT 10  
ID W64056 standard; Protein; 501 AA.  
AC W64056; (first entry)  
DE Human IL-9 receptor protein variant #1.  
KW Interleukin-9 receptor; IL-9R; Asthma associated factor 2; AAF2;  
soluble; bronchial hyper-responsiveness; detection; diagnosis; human;  
eosinophil; HL60 cell; antagonist.  
OS Homo sapiens.  
PN WO9824904-A2.  
PD 11-JUN-1998.  
PF 02-DEC-1997; U21992.  
PR 01-DEC-1997; US-032224.  
PR 02-DEC-1996; US-032224.  
PA (MAGA-) MAGANIN PHARM INC.  
PI Grasso L, Holroyd KJ, Levitt RC, Nicolaides NC;  
WPI; 98-348150/30.  
PT Human interleukin-9 receptor variant (s) useful in treating asthma -  
and related disorders e.g. bronchial hyper-responsiveness and for  
diagnosing greater or alternatively less susceptibility to these  
conditions  
Disclosure; Page 1; 81pp; English.  
This sequence represents a human interleukin-9 (IL-9) receptor variant  
(also known as Asthma Associated Factor 2 or AAF2) which is used in the  
construction of protein variants which are soluble and can be  
administered to humans to alleviate asthma and related disorders e.g.  
bronchial hyper-responsiveness. The DNA molecules are useful to detect  
or diagnose susceptibility to such conditions. Cells used in this method  
may be e.g. eosinophils and HL60 cells. Cells lacking human IL-9 receptor  
can be transfected with the DNA and used to identify IL-9 pathway  
antagonists. The nucleic acids can also be used to design probes to  
detect other IL-9 receptor variants.  
Sequence 501 AA;

Query Match 6.7%; Score 175; DB 1; Length 501;  
Best Local Similarity 27.6%; Pred. No. 2.01e-05;  
Matches 50; Conservative 43; Mismatches 74; Indels 14; Gaps 12;  
Db 203 LRQMATELDDVBEERYTGWSE---WSQPVCFQAPQROGLIPPGWPGN-TLVAISI 258  
QY 183 FWRVKAMED-VYGPDTVPDSEVTCWQGE-IRDAETP-TPPKPKLSKFLISLSLA 239  
Db 259 FLLLTGPTYL-LFKLSPRVKRIFQNVPSAMFFQPLYSVHNGFQWMAHRAAGVLLSQ 317  
QY 240 ILLMVSLLLSLWKLW-RVKKFLIPSPDPKSIFFGLFEIHQGNFQEWITDTQNVHL-H 297  
Db 318 DCAGTPQGALEPCVQEAATLLTCGPAPKWSVALEEEQEGPG-T-RLPGN-LSSEDLVPA 374  
QY 298 KMGAEQES-GPEELVVLQAKTEASPRMLDPQTEKEASGSLQLPHQLOGGDVVTI 356  
Db 375 G 375  
QY 357 G 357  
RESULT 11  
ID R47517 standard; Protein; 507 AA.  
AC R47517;  
DE 24-JUN-1994 (first entry)  
DE MEL EPO receptor.  
KW Erythropoietin receptor; recombinant; murine; anaemia.  
OS Mus musculus.  
FH Key Location/Qualifiers  
FT Peptide 1..24  
FT protein /note= "signal"  
FT 25..507  
FT /note= "mature EPO receptor"  
FT modified\_site /note= "potential N-glycosylation site"  
FT 383  
FT modified\_site /note= "potential N-glycosylation site"  
FT 250..271  
FT region /note= "putative transmembrane region"  
FN US5278065-A.  
PN 11-JAN-1994.  
PD 03-FEB-1989; 306503.  
PR 03-FEB-1989; US-306503.  
PR 25-MAR-1991; US-678877.  
PA (CHIL-) CHILDRENS MEDICAL CENT.  
PA (GEM) GENETICS INST INC.  
PA (WHEE) WHITEHEAD INST BIOMEDICAL RES.  
PI D'andrea A, Jones SS, Wong GG;  
WPI; 94-025409/03.  
DR N-PSDB; Q53994.  
PT Recombinant DNA encoding erythropoietin receptor - used to  
develop prods. for study, treatment or diagnosis of disorders in  
which receptor is dysfunctional.  
PS Disclosure; Fig 2; 24pp; English.  
CC Mouse erythroleukaemia (MEL) cells were used to construct a cDNA  
library. The cDNA was used to transfect COS-1 cells and these were  
screened for radioiodinated erythropoietin (EPO) binding to isolate  
cDNA encoding the EPO receptor. The cDNA may be used to isolate the  
EPO receptor from other sources and to study treat or diagnose  
disorders in which the EPO receptor is dysfunctional. The EPO  
receptor may also be used to raise antibodies or for treating  
hypersensitivity to EPO or who have elevated levels of EPO. The prod.  
is pref. used for treating anaemias, primary proliferative polycythemia  
and secondary polycythemia.  
CC See also R47518.  
SQ Sequence 507 AA;  
Query Match 6.3%; Score 163; DB 1; Length 507;  
Best Local Similarity 46.8%; Pred. No. 2.02e-04;  
Matches 29; Conservative 9; Mismatches 20; Indels 4; Gaps 3;  
Db 247 LDFLILTLSL-ILVLISLLLVALLSHRRTLQKINWIPSPSEFEGFLTHKGNFQL 305

QY 229 LSKFILSLAILLMVSL--LSLWKLWRVKKFLI-PSVPDPKSIFFGLFEIHQGNFQE 285  
 Db 306 WL 307  
 QY 286 WI 287

RESULT 12  
 ID R69502 standard; Protein; 507 AA.  
 AC R69502;

DT 10-AUG-1995 (first entry)  
 DE Mouse erythropoietin receptor.  
 KW Erythropoietin receptor; anemia therapy; signal peptide;  
 OS transmembrane region; N-linked glycosylation.  
 SW Mus musculus.

FT Key Location/Qualifiers  
 FT peptide 1..24  
 FT protein /note= "signal peptide"  
 FT 25..507  
 FT /note= "mature protein"  
 FT modified\_site 75..77  
 FT domain /note= "N-linked glycosylation site"  
 FT 250..271  
 FT /note= "transmembrane region"  
 FT modified\_site 383..385  
 FT /note= "N-linked glycosylation site"

US5378808-A.  
 03-JAN-1995.  
 PD 03-FEB-1989; 306503.  
 PF 03-FEB-1989; US-306503.  
 PR 25-MAR-1991; US-678877.  
 PR 10-JUN-1993; US-075069.  
 PA (GENE ) GENETICS INST INC.  
 PI D'andrea A, Jones SS, Wong GG;  
 DR WPI; 95-051310/07.  
 DR N-PSDB; Q81891.

PT New recombinant erythropoietin receptor polypeptide(s) - used for  
 PT detection, purificn. and therapy and for prodn. of antibodies for  
 PT detection and therapy  
 PS Claim 1; Fig 2; 24pp; English.  
 CC The sequence corresponds to a mouse erythropoietin receptor,  
 CC including putative signal peptide and transmembrane regions, and 2  
 CC N-linked glycosylation sites. The protein is derived from mouse  
 CC erythroleukemia cells and may be used in drug design or in  
 CC pharmaceutical compositions for therapy of anemia.  
 SQ Sequence 507 AA;

Query Match 6.3%; Score 163; DB 1; Length 507;  
 Best Local Similarity 46.8%; Pred. No. 2.02e-04;  
 Matches 29; Conservative 9; Mismatches 20; Indels 4; Gaps 3;

Db .247 LDPLILTL-ILVLISLLTLVALLSHRRTLQOKIWPICSPSEFEGFLFTHKGNFOL 305  
 QY 229 LSKFILSLAILLMVSL--LSLWKLWRVKKFLI-PSVPDPKSIFFGLFEIHQGNFQE 285  
 Db 306 WL 307  
 QY 286 WI 287

RESULT 13  
 ID R50327 standard; Protein; 507 AA.  
 AC R50327;

DT 19-OCT-1994 (first entry)  
 DE Mouse soluble EPO receptor protein.  
 KW Murine; soluble; erythropoietin; EPO; receptor protein; SEPO-R; drug;  
 OS antigen; diagnostic agent; biochemical reagent.  
 SW Mus musculus.

FT Key Location/Qualifiers  
 FT modified\_site 75..77  
 FT /note= "N-linked glycosylation site"  
 PN J06038787-A.

PD 15-FEB-1994.  
 PF 04-MAR-1992; 082865.  
 PR 04-MAR-1992; JP-082865.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 DR WPI; 94-094847/12.  
 DR N-PSDB; Q44854.  
 FT Soluble erythropoietin receptor protein - and DNA coding for  
 PT SEPO-R, useful as diagnostic reagent  
 PS Disclosure; Fig 1; 9pp; Japanese.  
 CC This sequence represents the murine soluble erythropoietin (EPO)  
 CC receptor protein (SEPO-R). This protein is able to bind to EPO and  
 CC has antigenicity as an EPO receptor. The molecular weight of the  
 CC full length protein is pref 33 or 29 kd. The protein is useful as a  
 CC drug, as a diagnostic agent and a biochemical reagent.  
 SQ Sequence 507 AA;

Query Match 6.3%; Score 163; DB 1; Length 507;  
 Best Local Similarity 46.8%; Pred. No. 2.02e-04;  
 Matches 29; Conservative 9; Mismatches 20; Indels 4; Gaps 3;

Db 247 LDPLILTL-ILVLISLLTLVALLSHRRTLQOKIWPICSPSEFEGFLFTHKGNFOL 305  
 QY 229 LSKFILSLAILLMVSL--LSLWKLWRVKKFLI-PSVPDPKSIFFGLFEIHQGNFQE 285  
 Db 306 WL 307  
 QY 286 WI 287

RESULT 14

ID R06511 standard; protein; 507 AA.  
 AC R06511;  
 DT 04-JAN-1991 (first entry)  
 DE EPO receptor sequence deduced from DNA of clone 190.  
 KW Erythropoietin; Diamond Blackfan anaemia; polycythemia vera.  
 OS Mus musculus.

FT Key Location/Qualifiers  
 FT peptide 1..24  
 FT domain /label=signal peptide  
 FT 25..248  
 FT /label=extracellular domain  
 FT domain /note=EPO binding region  
 FT 248..271  
 FT domain /label=transmembrane domain  
 FT 272..507  
 FT /label=intracellular domain  
 FT modified\_site 75..77  
 FT /label=N-linked\_glycos  
 FT modified\_site 182..184  
 FT /label=N-linked\_glycos

WO9008822-A.

PD 09-AUG-1990.  
 PF 01-FEB-1990; U00635.  
 PR 03-FEB-1989; US-306503.  
 PA (GENE-) GENETICS INST INC.  
 PI (WHIT-) WHITEHEAD INST.  
 PI D'andrea A, Wong G;  
 DR WPI; 90-260931/34.  
 DR N-PSDB; Q05747.

PT Erythropoietin receptor and gene - used for developing reagents  
 PT and systems to control and study erythropoiesis.  
 PS Disclosure; Fig 1; 53pp; English.  
 CC The sequence was deduced from DNA from a clone isolated from a  
 CC cDNA library prep. from uninduced murine erythroleukemia cells.  
 CC It is a type I transmembrane protein with binding affinity for EPO.  
 CC The gene and recombinant EPO receptor produced on expression of  
 CC the DNA are used to develop reagents and systems to control and  
 CC study erythropoiesis. It is believed that the EPO receptor is  
 CC dysfunctional in individuals with Diamond Blackfan anaemia, and  
 CC may be hyperactive in polycythemia vera.  
 CC See also R06512 (human EPO receptor).  
 SQ Sequence 507 AA;





US-09-376-430-2-02.ra1

Thu May 11 06:49:26 2000

\*\*\*\*\*

MPSEARCH

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:18:35 2000; Maspar time 45.11 Seconds  
Tabular output not generated.

Title: >US-09-376-430-2  
Description: (23-371) from US09376430A.pep (2 of 25)  
Perfect Score: 2601  
Sequence: 1 QGAAEGVQIIIFNLETV.....DVTIGTFVNDRSYVAL 349

Scoring table: PAM 150  
Gap 11

Searched: 131253 seqs, 12956647 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-issued  
1:5A\_COMB 2:5B\_COMB 3:PCT\_COMB 4:backfiles1

Statistics: Mean 32.722; Variance 150.110; scale 0.218

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	ID	Description	Pred. No.
1	195	7.5	230	1	US-08-595-Sequence 11, Applicati	5.53e-07
2	195	7.5	230	1	US-08-052-Sequence 11, Applicati	5.53e-07
3	195	7.5	230	1	US-08-595-Sequence 9, Applicati	5.53e-07
4	195	7.5	230	1	US-08-052-Sequence 9, Applicati	5.53e-07
5	195	7.5	347	1	US-08-595-Sequence 7, Applicati	5.53e-07
6	195	7.5	347	1	US-08-052-Sequence 7, Applicati	5.53e-07
7	195	7.5	369	1	US-08-595-Sequence 4, Applicati	5.53e-07
8	195	7.5	369	1	US-08-052-Sequence 4, Applicati	5.53e-07
9	191	7.3	536	1	US-08-456-Sequence 12, Applicati	1.17e-06
10	191	7.3	536	1	US-08-164-Sequence 12, Applicati	1.17e-06
11	188	7.2	369	2	Sequence 2, Applicati	2.04e-06
12	188	7.2	369	3	PCT-US94-0Sequence 69, Applicati	2.04e-06
13	178	6.8	522	2	US-08-456-Sequence 10, Applicati	1.31e-05
14	178	6.8	522	2	US-08-164-Sequence 10, Applicati	1.31e-05
15	166	6.4	379	2	US-08-456-Sequence 8, Applicati	1.18e-04
16	166	6.4	379	1	US-08-164-Sequence 8, Applicati	1.18e-04
17	166	6.4	468	2	US-08-456-Sequence 7, Applicati	1.18e-04
18	166	6.4	468	1	US-08-164-Sequence 7, Applicati	1.18e-04
19	160	6.2	508	2	US-08-850-Sequence 10, Applicati	5.03e-04
20	158	6.1	258	1	US-08-336-Sequence 5, Applicati	1.27e+00
21	130	5.0	383	1	US-08-609-Sequence 2, Applicati	1.27e+00
22	113	4.3	897	1	US-07-980-Sequence 2, Applicati	2.09e+00
23	110	4.2	551	4	5198359-2Patent No. 5198359.	

24 110 4.2 551 4 5449756-2 Patent No. 5449756. 2.09e+00  
25 107 4.1 459 4 5194375-2 Patent No. 5194375. 3.42e+00  
26 100 3.8 788 2 US-07-728-Sequence 27, Applicati 1.06e+01  
27 100 3.8 913 1 US-08-445-Sequence 4, Applicatio 1.06e+01  
28 100 3.8 919 1 US-08-336-Sequence 2, Applicatio 1.06e+01  
29 99 3.8 982 2 US-08-673-Sequence 4, Applicatio 1.25e+01  
30 99 3.8 1073 1 US-08-241-Sequence 1, Applicatio 1.25e+01  
31 99 3.8 1073 1 US-07-695-Sequence 1, Applicatio 1.25e+01  
32 99 3.8 1091 1 US-07-695-Sequence 3, Applicatio 1.25e+01  
33 99 3.8 1091 1 US-08-241-Sequence 3, Applicatio 1.25e+01  
34 95 3.7 141 1 US-07-695-Sequence 5, Applicatio 2.35e+01  
35 95 3.7 141 1 US-08-241-Sequence 5, Applicatio 2.35e+01  
36 95 3.7 149 1 US-08-241-Sequence 7, Applicatio 2.35e+01  
37 95 3.7 149 1 US-07-695-Sequence 7, Applicatio 2.35e+01  
38 96 3.7 505 1 US-08-149-Sequence 16, Applicati 2.00e+01  
39 96 3.7 505 1 US-08-317-Sequence 16, Applicati 2.00e+01  
40 97 3.7 539 4 5449756-4 Patent No. 5449756. 1.71e+01  
41 97 3.7 539 4 5198359-4 Patent No. 5198359. 2.00e+01  
42 96 3.7 799 1 US-08-054-Sequence 2, Applicatio 2.00e+01  
43 96 3.7 822 2 US-08-474-Sequence 9, Applicatio 2.00e+01  
44 96 3.7 822 2 US-08-474-Sequence 9, Applicatio 2.00e+01  
45 96 3.7 822 2 US-08-472-Sequence 8, Applicatio 2.00e+01

ALIGNMENTS

RESULT 1  
ID US-08-595-974-11 STANDARD; PRT; 230 AA.

XX xxxxxx

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 11, Application US/08595974

Sequence 11, Application US/08595974

Patent No. 5705608

GENERAL INFORMATION:

APPLICANT: SUGAMURA, KAZUO

APPLICANT: TAKESHITA, TOSHIKAZU

APPLICANT: ASAO, HIRONOBU

APPLICANT: NAKAMURA, MASATAKA

APPLICANT: SHIMAMURA, TOSHIRO

APPLICANT: SUZUKI, MANABU

APPLICANT: HAMURO, JUNJI

TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: P.C. SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

STREET: 1755 S. Jefferson Davis Highway, Suite 400

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM: disk

MEDIUM TYPE: Floppy

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/595,974

FILING DATE: 06-FEB-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/052,205

FILING DATE: 22-APR-1993

APPLICATION NUMBER: JP 104947/1992

FILING DATE: 23-APR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Obion, No. 5705608man F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 10-615-0X



US-09-376-430-2-02.rai

Thu May 11 06:49:26 2000

CC APPLICATION NUMBER: US/08/595,974  
 CC FILING DATE: 06-FEB-1996  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/052,205  
 CC FILING DATE: 22-APR-1993  
 CC APPLICATION NUMBER: JP 104947/1992  
 CC FILING DATE: 23-APR-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: OBLON, NO. 5705608man F.  
 CC REGISTRATION NUMBER: 24,618  
 CC REFERENCE/DOCKET NUMBER: 10-615-0X  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (703) 413-3000  
 CC TELEFAX: (703) 413-2220  
 CC TELEX: 248855 OPAT UR  
 CC INFORMATION FOR SEQ ID NO: 9:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 252 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 252 AA; 29505 MW; 347903 CN;  
 CC  
 CC Query Match 7.5%; Score 195; DB 1; Length 252;  
 CC Best Local Similarity 25.9%; Pred. No. 5.53e-07;  
 CC Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 CC  
 Db 59 EVOCFVFNVMNCTNSSEPOPTNLTHYWKNSDNKVKCSHYLSEITSCGQLQ 118  
 Qy 31 QIQIIVFNLTVQVTWNSKYSR-TNLTFFHYRF-NGD-EAYDOCTNYLLQEGHTSGCLLD 87  
 Db 119 KKEIHLVOTFVVOLODPREPRQATQMLKQLNLPWAPENLTLHLKLSQLELNWNR 178  
 Qy 88 AEQRDDILYFSIR-NGTH-PVFTASRWV-VYVL-KPSSPKHVRF-SWHQDAVTVCSDLS 142  
 Db 179 LNHCHLHVQYRTDMDHSTEQSVYRHKFSLPSVDGQKRYTFVRVRSR-FNPLCGSAQHW 237  
 Qy 143 YGD-LLYEVQYRSPFDTEW-QSKQENTCNVTIEGLDAEKCYSFVVRKAMEDVYGPDTYP 200  
 Db 238 SEWSHPHWH 246  
 Qy 201 SDWSEVTCW 209  
 CC  
 CC RESULT 4  
 CC ID US-08-052-205-9 STANDARD; PRT; 252 AA.  
 CC XX  
 CC AC xxxxxx  
 CC DT  
 CC DE  
 CC XX  
 CC Sequence 9, Application US/08052205  
 CC Patent No. 5510259  
 CC GENERAL INFORMATION:  
 CC APPLICANT: SUGAMURA, KAZUO  
 CC APPLICANT: TAKESHITA, TOSHIKAZU  
 CC APPLICANT: ASAO, HIRONOBU  
 CC APPLICANT: NAKAMURA, MASATAKA  
 CC APPLICANT: SHIMAMURA, TOSHIRO  
 CC APPLICANT: SUZUKI, MANABU  
 CC APPLICANT: HAMURO, JUNJI  
 CC TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE  
 CC NUMBER OF SEQUENCES: 21  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 CC STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 CC CITY: Arlington  
 CC STATE: Virginia  
 CC COUNTRY: U.S.A.  
 CC

CC ZIP: 22202  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/052,205  
 CC FILING DATE: 19930422  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: JP 104947/1992  
 CC FILING DATE: 23-APR-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: OBLON, NO. 5510259man F.  
 CC REGISTRATION NUMBER: 24,618  
 CC REFERENCE/DOCKET NUMBER: 10-615-0X  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (703) 413-3000  
 CC TELEFAX: (703) 413-2220  
 CC TELEX: 248855 OPAT UR  
 CC INFORMATION FOR SEQ ID NO: 9:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 252 amino acids  
 CC TYPE: AMINO ACID  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 252 AA; 29505 MW; 347903 CN;  
 CC  
 CC Query Match 7.5%; Score 195; DB 1; Length 252;  
 CC Best Local Similarity 25.9%; Pred. No. 5.53e-07;  
 CC Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 CC  
 Db 59 EVOCFVFNVMNCTNSSEPOPTNLTHYWKNSDNKVKCSHYLSEITSCGQLQ 118  
 Qy 31 QIQIIVFNLTVQVTWNSKYSR-TNLTFFHYRF-NGD-EAYDOCTNYLLQEGHTSGCLLD 87  
 Db 119 KKEIHLVOTFVVOLODPREPRQATQMLKQLNLPWAPENLTLHLKLSQLELNWNR 178  
 Qy 88 AEQRDDILYFSIR-NGTH-PVFTASRWV-VYVL-KPSSPKHVRF-SWHQDAVTVCSDLS 142  
 Db 179 LNHCHLHVQYRTDMDHSTEQSVYRHKFSLPSVDGQKRYTFVRVRSR-FNPLCGSAQHW 237  
 Qy 143 YGD-LLYEVQYRSPFDTEW-QSKQENTCNVTIEGLDAEKCYSFVVRKAMEDVYGPDTYP 200  
 Db 238 SEWSHPHWH 246  
 Qy 201 SDWSEVTCW 209  
 CC  
 CC RESULT 5  
 CC ID US-08-595-974-7 STANDARD; PRT; 347 AA.  
 CC XX  
 CC AC xxxxxx  
 CC DT  
 CC DE  
 CC XX  
 CC Sequence 7, Application US/08595974  
 CC Patent No. 5705608  
 CC GENERAL INFORMATION:  
 CC APPLICANT: SUGAMURA, KAZUO  
 CC APPLICANT: TAKESHITA, TOSHIKAZU  
 CC APPLICANT: ASAO, HIRONOBU  
 CC APPLICANT: NAKAMURA, MASATAKA  
 CC APPLICANT: SHIMAMURA, TOSHIRO  
 CC APPLICANT: SUZUKI, MANABU  
 CC APPLICANT: HAMURO, JUNJI  
 CC TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE  
 CC NUMBER OF SEQUENCES: 21  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 CC

CC ADDRESSEE: P.C.  
 CC STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 CC CITY: Arlington  
 CC STATE: Virginia  
 CC COUNTRY: U.S.A.  
 CC ZIP: 22202

## COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/595,974  
 CC FILING DATE: 06-FEB-1996  
 CC CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/052,205  
 CC FILING DATE: 22-APR-1993  
 CC APPLICATION NUMBER: JP 104947/1992  
 CC FILING DATE: 23-APR-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Oblon, No. 5705608man F.  
 CC REGISTRATION NUMBER: 24,618  
 CC REFERENCE/DOCKET NUMBER: 10-615-0X  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (703) 413-3000  
 CC TELEFAX: (703) 413-2220  
 CC TELEFAX: (703) 413-3000

## INFORMATION FOR SEQ ID NO: 7:

CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 347 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 347 AA; 39920 MW; 666553 CN;

Query Match 7.5%; Score 195; DB 1; Length 347;  
 Best Local Similarity 25.9%; Pred. No. 5.53e-07;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;

Db 37 EVQCFVNFVNMCTNSSEPOPTNLTHYWKNSDNDKVKCSHYLSEITSGCOLQ 96  
 QY 31 QIQIYFNLEIVQVTVNASKYSR-TNLTTHYRF-NGD-EAYDOCTNYLLOEGHTSGCLLD 87  
 Db 97 KKEHLYQTFVVQLODPREPQRQATQMLKQNLVIPWAPENLTLHKLSELSOLELNNRNF 156  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYYL-KPSSPKHVRF-SWHQDAVTVTCSDL 142  
 Db 157 LNHCHLHVQYRTDWDHSHWTEQSDYRHKFSLPSVDGOKRYTFVRYSR-FNPLCGSAQHW 215  
 QY 143 YGD-LLYEVQYRSFPDTEW-QSKQENTCNVTIEGLDAEKCYSFWRVVRKAMEDVYGPDTYP 200  
 Db 216 SEMSHPIHW 224  
 QY 201 SDWSEVTCW 209

RESULT 6  
 ID US-08-052-205-7 STANDARD; PRT; 347 AA.  
 XX  
 AC xxxxxx  
 XX  
 DT

Sequence 7, Application US/08052205

Sequence 7, Application US/08052205  
 Patent No. 5510259

## GENERAL INFORMATION:

CC APPLICANT: SUGAMURA, KAZUO  
 CC APPLICANT: TAKESHITA, TOSHIKAZU  
 CC APPLICANT: ASAO, HIRONOBU  
 CC APPLICANT: NAKAMURA, MASATAKA

CC APPLICANT: SHIMAMURA, TOSHIRO  
 CC APPLICANT: SUZUKI, MANABU  
 CC APPLICANT: HAMURO, JUNJI  
 CC TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE  
 CC NUMBER OF SEQUENCES: 21  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 CC ADDRESSEE: P.C.  
 CC STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 CC CITY: Arlington  
 CC STATE: Virginia  
 CC COUNTRY: U.S.A.  
 CC ZIP: 22202

## COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/052,205  
 CC FILING DATE: 19930422  
 CC CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: JP 104947/1992  
 CC FILING DATE: 23-APR-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Oblon, No. 5510259man F.  
 CC REGISTRATION NUMBER: 24,618  
 CC REFERENCE/DOCKET NUMBER: 10-615-0X  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (703) 413-3000  
 CC TELEFAX: (703) 413-2220  
 CC TELEFAX: 248855 OPAT UR

## INFORMATION FOR SEQ ID NO: 7:

CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 347 amino acids  
 CC TYPE: AMINO ACID  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 347 AA; 39920 MW; 666553 CN;

Query Match 7.5%; Score 195; DB 1; Length 347;  
 Best Local Similarity 25.9%; Pred. No. 5.53e-07;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;

Db 37 EVQCFVNFVNMCTNSSEPOPTNLTHYWKNSDNDKVKCSHYLSEITSGCOLQ 96  
 QY 31 QIQIYFNLEIVQVTVNASKYSR-TNLTTHYRF-NGD-EAYDOCTNYLLOEGHTSGCLLD 87  
 Db 97 KKEHLYQTFVVQLODPREPQRQATQMLKQNLVIPWAPENLTLHKLSELSOLELNNRNF 156  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYYL-KPSSPKHVRF-SWHQDAVTVTCSDL 142  
 Db 157 LNHCHLHVQYRTDWDHSHWTEQSDYRHKFSLPSVDGOKRYTFVRYSR-FNPLCGSAQHW 215  
 QY 143 YGD-LLYEVQYRSFPDTEW-QSKQENTCNVTIEGLDAEKCYSFWRVVRKAMEDVYGPDTYP 200  
 Db 216 SEMSHPIHW 224  
 QY 201 SDWSEVTCW 209

RESULT 7  
 ID US-08-595-974-4 STANDARD; PRT; 369 AA.  
 XX  
 AC xxxxxx  
 XX  
 DT

Sequence 4, Application US/08595974

Sequence 4, Application US/08595974  
 Patent No. 5705608



RESULT 8  
ID US-08-052-205-4  
XX  
AC xxxxxx  
PRT; 369 AA.

```

DB* 238 SEWSHP1HW 240
      | : | | |
OV 201 SDWSEVTCW 209

```



US-09-376-430-2-02.ra1

Thu May 11 06:49:26 2000

```

XX Sequence 2, Application US/08424224
DE
XX Sequence 2, Application US/08424224
XX Patent No. 5912173
CC GENERAL INFORMATION:
CC APPLICANT: LEONARD, WARREN J. CDNA AND
CC TITLE OF INVENTION: MURINE IL-2R
CC NUMBER OF INVENTIONS: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: MORGAN & FINNEGAN
CC STREET: 345 PARK AVE.
CC CITY: NEW YORK
CC STATE: NEW YORK
CC COUNTRY: USA
CC ZIP: 10154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: FLOPPY DISK
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: WORD PERFECT # 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/424,224
CC FILING DATE:
CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/121,435
CC FILING DATE: 14-SEPT-1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: WILLIAM S. FEILER
CC REGISTRATION NUMBER: 26,728
CC REFERENCE/DOCKET NUMBER: 2026-4061US1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212-758-4800
CC TELEFAX: 212-751-6849
CC TELEX: 421792
CC INFORMATION FOR SEQ ID NO: 2:
CC INDIVIDUAL ISOLATE: IL-2R
CC SEQUENCE 369 AA; 42241 MW; 762797 CN;

Query Match 7.2%; Score 188; DB 2; Length 369;
Best Local Similarity 26.8%; Pred. No. 2,04e-06;
Matches 51; Conservative 52; Mismatches 75; Indels 12; Gaps 11;

Db 59 EVQCFVFNIEYMNCTWNSSEPOATNLTLYRYKVSNNFTQECSHYLFSEKTSQCQIQ 118
QY 31 QIQIIFNLETVQVWTWNAKYSR-TNLTIFYRFN-GDE-AYDQCTNLLQEGHTSGCLLD 87

Db 119 KEDIQYQTFVVOQDPQKORRAVQKLNQNLVPRAPENLTLSNLSOLELRWKSRRH 178
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRWV-VYIL-KPSSPKHVRFS-WHQDQAVTV--TCSD 140

Db 179 IKERCLOYLYQVRSNRDRSWTELIVNHEPRFSLPSVDLKRKYTFVRVRSRY-NPICGSSQ 237
QY 141 LSYGDLLEYQVRSFDFIEM-QSKQCNCTNVTIEGLDAEKCYSFVWRVKAMEDVIGPDY 199

Db 238 NKSQSPVHW 247
QY 200 PSDWSEVTCW 209

RESULT 12
ID PCT-US94-02891-69 STANDARD: PRT; 369 AA.

```

```

XX
AC
XX
DT
XX
DE
XX
XX
Sequence 69, Application PC/TUS9402891
Sequence 69, Application PC/TUS9402891
GENERAL INFORMATION:
APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
APPLICANT: REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN
APPLICANT: SERVICES
APPLICANT: OFFICE OF TECHNOLOGY TRANSFER, NATIONAL
APPLICANT: INSTITUTES OF HEALTH, BOX OTT, BETHESDA, MARYLAND 20892 USA
TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: XSCID
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVE.
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: FLOPPY DISK
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: WORD PERFECT # 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/02891
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA: 08/031,143
CC APPLICATION NUMBER: 12-MAR-1993
CC FILING DATE: 08/121,435
CC APPLICATION NUMBER: 14-SEPT-1993
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: WILLIAM S. FEILER
CC REGISTRATION NUMBER: 26,728
CC REFERENCE/DOCKET NUMBER: 2026-4061
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212-758-4800
CC TELEFAX: 212-751-6849
CC TELEX: 421792
CC INFORMATION FOR SEQ ID NO: 69:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 369
CC TYPE: AMINO ACID
CC TOPOLOGY: UNKNOWN
CC MOLECULE TYPE: PROTEIN
CC DESCRIPTION: NO
CC HYPOTHETICAL: NO
CC ORIGINAL SOURCE:
CC ORGANISM: MURINE
CC INDIVIDUAL ISOLATE: IL-2R
CC SEQUENCE 369 AA; 42241 MW; 762797 CN;

Query Match 7.2%; Score 188; DB 3; Length 369;
Best Local Similarity 26.8%; Pred. No. 2,04e-06;
Matches 51; Conservative 52; Mismatches 75; Indels 12; Gaps 11;

Db 59 EVQCFVFNIEYMNCTWNSSEPOATNLTLYRYKVSNNFTQECSHYLFSEKTSQCQIQ 118
QY 31 QIQIIFNLETVQVWTWNAKYSR-TNLTIFYRFN-GDE-AYDQCTNLLQEGHTSGCLLD 87

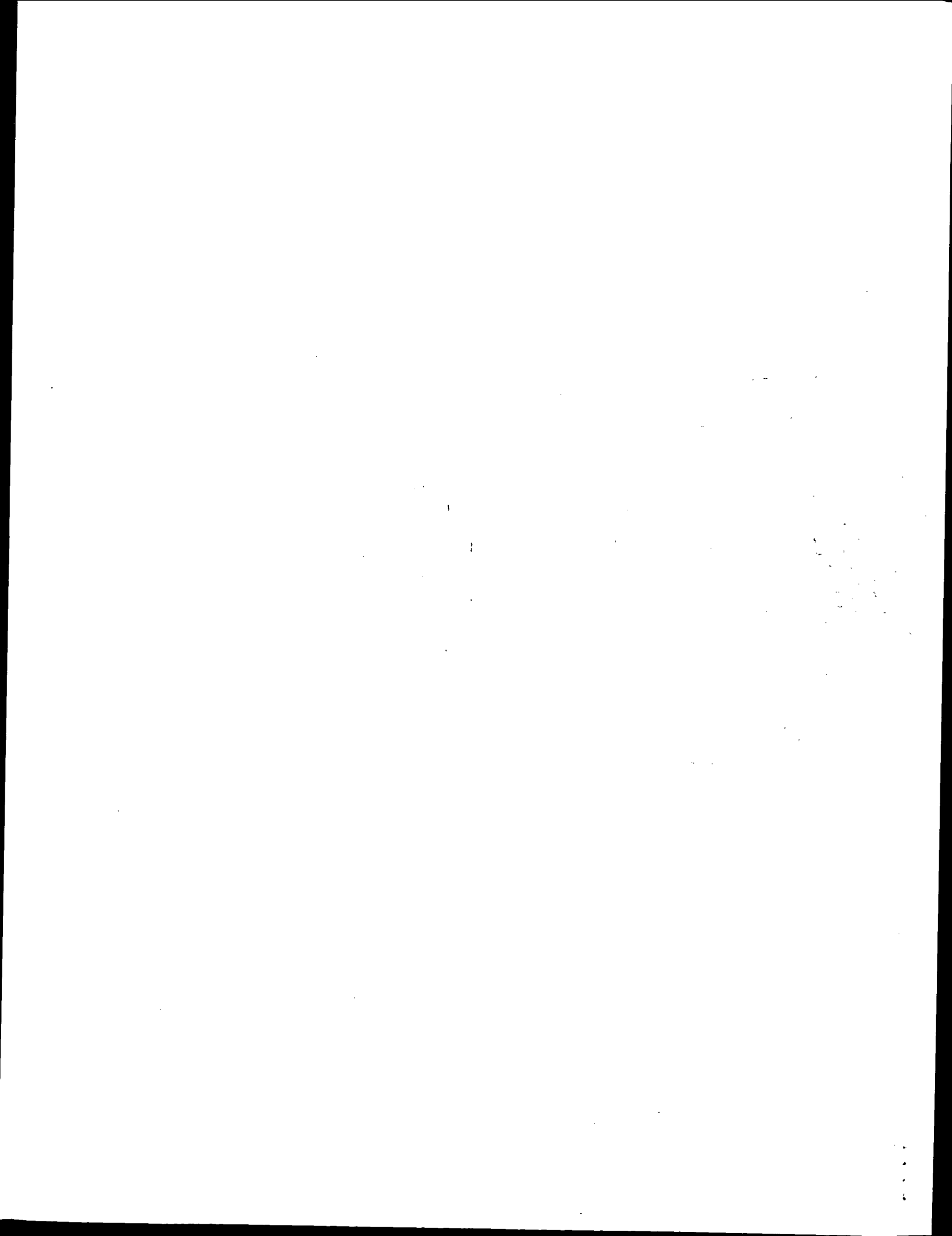
Db 119 KEDIQYQTFVVOQDPQKORRAVQKLNQNLVPRAPENLTLSNLSOLELRWKSRRH 178
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRWV-VYIL-KPSSPKHVRFS-WHQDQAVTV--TCSD 140

Db 179 IKERCLOYLYQVRSNRDRSWTELIVNHEPRFSLPSVDLKRKYTFVRVRSRY-NPICGSSQ 237

```







US-09-376-430-2-02.rpr

Thu May 11.06:49:27 2000

\*\*\*\*\*  
M P E S R L H  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd  
MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:17:14 2000; MasPar time 23.33 Seconds  
Tabular output not generated.  
705.727 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (23-371) from US09376430A.pap (2 of 25)  
Perfect Score: 2601  
Sequence: 1 QGGAAGVQIIYFNLETV.....DVTIGGTFVMDRSYVAL 349  
Scoring table: PAM 150  
Gap 11  
Searched: 142080 seqs, 47172406 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: pir62  
1:pir1 2:pir2 3:pir3 4:pir4  
Statistics: Mean 47.775; Variance 94.695; scale 0.505

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES				Pred. No.	
Result No.	Score	Query Match Length DB ID	Description	Pred. No.	
1	210	8.1	interleukin-2 recepto	3.00e-17	
2	195	7.5	interleukin-2 recepto	1.07e-14	
3	188	7.2	interleukin-2 recepto	1.59e-13	
4	178	6.8	gene ffr-2 protein	7.11e-12	
5	176	6.8	interleukin-9 recepto	1.51e-11	
6	166	6.4	interleukin-9 recepto	6.18e-10	
7	164	6.3	erythropoietin recept	1.29e-09	
8	163	6.3	erythropoietin recept	1.86e-09	
9	160	6.2	erythropoietin recept	5.53e-09	
10	115	4.4	interleukin-3 recepto	4.22e-02	
11	113	4.3	cytokine receptor com	1.05e-01	
12	110	4.2	probable yop transloc	1.41e-01	
13	109	4.2	probable aspartate am	1.90e-01	
14	108	4.2	cytochrome P450 2C7	1.05e-01	
15	110	4.2	interleukin-2 recepto	1.05e-01	
16	110	4.2	somatotropin receptor	1.05e-01	
17	110	4.2	somatotropin receptor	1.05e-01	
18	110	4.2	somatotropin receptor	1.05e-01	
19	110	4.2	somatotropin receptor	1.05e-01	
20	108	4.2	cytokine receptor com	1.90e-01	
21	106	4.1	integrin alpha-7C cha	3.43e-01	
22	107	4.1	interleukin-7 recepto	2.56e-01	
23	107	4.1	interleukin-7 recepto	2.56e-01	

24 106 4.1 490 2 A36122 cytochrome P450 2C13 3.43e-01  
25 106 4.1 1073 1 OXHUXH heat-stable enterotox 3.43e-01  
26 106 4.1 1106 2 S38783 integrin alpha chain 3.43e-01  
27 105 4.0 378 2 S00842 leukosialin precursor 4.60e-01  
28 105 4.0 1135 2 I61186 alpha-7 integrin - mo 4.60e-01  
29 104 4.0 1151 2 A45226 integrin alpha-1 subu 6.14e-01  
30 104 4.0 1180 2 A35854 integrin alpha-1 chai 1.45e+00  
31 101 3.9 184 2 H72248 ribosomal protein L5 1.45e+00  
32 101 3.9 359 2 I49341 MIP-1 alpha receptor 1.45e+00  
33 101 3.9 381 2 S24611 latent membrane prote 1.45e+00  
34 102 3.9 386 1 Q0BE50 latent membrane prote 1.45e+00  
35 101 3.9 404 1 L0BECA latent membrane prote 1.45e+00  
36 101 3.9 418 2 D70038 maltodextrin transpor 1.45e+00  
37 102 3.9 443 2 G75038 probable na+/h+ antip 1.09e+00  
38 102 3.9 861 2 S77086 hypothetical protein 1.09e+00  
39 102 3.9 1062 2 JC5951 integrin alpha 7 chai 1.09e+00  
40 102 3.9 1137 2 JC5950 integrin alpha 7 chai 1.09e+00  
41 100 3.8 169 2 S76289 hypothetical protein 1.92e+00  
42 100 3.8 220 2 I80329 receptor tyrosine kin 1.92e+00  
43 100 3.8 669 2 D64137 beta7 protein homolog 1.92e+00  
44 100 3.8 876 2 A48508 protein-tyrosine kina 1.92e+00  
45 100 3.8 1021 2 I39207 leukocyte surface pro 1.92e+00

ALIGNMENTS  
RESULT 1  
ENTRY #type complete  
TITLE interleukin-2 receptor gamma chain precursor - dog  
ORGANISM #formal\_name Canis lupus familiaris #common\_name dog  
DATE 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 23-Jul-1999  
ACCESSIONS A55718  
REFERENCE #authors Henthorn, P.S.; Somberg, R.L.; Fimiani, V.M.; Puck, J.M.; Patterson, D.F.; Felsburg, P.J.  
#journal Genomics (1994) 23:69-74  
#title IL-28gamma gene microdeletion demonstrates that canine X-linked severe combined immunodeficiency is a homologue of the human disease.  
#cross-references MIM:95130114  
#accession A55718  
#status Preliminary  
#molecule\_type mRNA  
##residues 1-373 #label HEN  
##cross-references GB:U04361; NID:9517411; PIDN:AAC48403.1; PID:9517412  
CLASSIFICATION #superfamily interleukin-2 receptor gamma chain  
KEYWORDS cytokine receptor; duplication  
SUMMARY #length 373 #molecular-weight 42516 #checksum 7960  
Query Match 8.1%; Score 210; DB 2; Length 373;  
Best Local Similarity 25.9%; Pred. No. 3.00e-17;  
Matches 49; Conservative 56; Mismatches 73; Indels 11; Gaps 10;

Db 59 EVQCFVFNVEYMNCTWNSSEPRPTNLTHYWKNSNDKVOEGCHYLFREVTAGCWLQ 118  
QY 31 QIQIIVFNLETVQVTWNAKYSR-TNLTFRYF-NGDEA-YDQCTNYLLQEGHTSGCLLD 87  
Db 119 KEETHLYETVVDLPDRPREPROSTQKLQNLVWPAPENLTHLNLSQLELSWSNRH 178  
QY 88 AEQRD--DILYFSIRNGTHPVFTAS-RWVVYL-KPSSPKHVR-F-SWHDQAVTVTCSDL 142  
Db 179 LDHGLEHVQVRSMDRSWTSQSDVDRNSTSLPSVDGKFTFRVGRY-NPLCGSAQRW 237  
QY 143 YGD-LLEVEQVRSFDEW-QSKENCTNVTIEGLDAEKCYSFWRVKAMEDVYGPDTYP 200  
Db 238 SEWSHPHWH 246  
QY 201 SDWSEYTCW 209

RESULT 2  
ENTRY #type complete  
A42565





```
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 1-369 #label RES
##cross-references GB:S75852; NID:9861554; PID:AAB32904.1; PID:9861555
GENETICS
#gene
#introns
#complex
The high affinity receptor is a heterotrimer of alpha (see
PIR:UHM52), beta (see PIR:A35052), and gamma chains;
heterodimers of alpha or beta and gamma chains are
intermediate affinity receptors.
FUNCTION
#description
#pathway
interleukin-2 stimulated growth and differentiation of T
cells, B cells, NK cells, LAK cells, monocytes,
macrophages, and oligodendrocytes
#superfamily interleukin-2 receptor gamma chain
#cytokine_receptor; duplication; glycoprotein; transmembrane
protein
FEATURE
1-22 #domain signal sequence #status predicted #label SIG\
23-369 #product interleukin-2 receptor gamma chain #status
predicted #label MAN\
256-284 #domain transmembrane #status predicted #label TM\
71,75,84,96,159, #binding_site carbohydrate (Asn) (covalent) #status
predicted
164,306 #length 369 #molecular_weight 42241 #checksum 6734
SUMMARY
Query Match 7.2%; Score 188; DB 2; Length 369;
Best Local Similarity 26.8%; Pred. No. 1.59e-13;
Matches 51; Conservative 52; Mismatches 75; Indels 12; Gaps 11;
Db 59 EVOCVENIEYVNWSSSEPAQNLTHYRYKYSDNNFOECSHYLFSEKITSQCQIQ 118
QY 31 QIOIIEVLETVQVWNASKYSR-TNLTFHYRFN-GDE-AYDQCTNYLLQEGHTSGCLLD 87
Db 119 KEDILQYOTFVQLODPQKQPRQVQKLNQNLQVPRAPENLTLSNLSOLELRWKS 178
QY 88 AQRODILYFSIR-NGTH-PVFTASRWV-VYVL-KPSSPKHVRFS-WHODAVTV--TCS 140
Db 179 IKERQYLYQVRSNDRSWEITLVNHEPRLSPVDELKRYTFVRVRSY-NPIGSSQ 237
QY 141 LSYGDLLEYQVRSFPTDTEW-QSKQENTCNVTIEGLDAEKCYSFVVRKAMEDVYGP 199
Db 238 WSKSQPVHW 247
QY 200 PSDWSEVTCW 209
RESULT 4
ENTRY
TITLE
ORGANISM
DATE
#type complete
#gene gfi-2 protein - rat
#formal_name Rattus norvegicus #common_name Norway rat
#sequence_revision 26-Jul-1996 #text_change
28-Feb-1997
ACCESSIONS
REFERENCE
#authors
#journal
#title
#cross-references MUID:95055995
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-467 #label RES
#cross-references GB:L36453; NID:9598371; PID:9598372
GENETICS
#gene
#introns
#complex
The high affinity receptor is a heterotrimer of alpha (see
PIR:UHM52), beta (see PIR:A35052), and gamma chains;
heterodimers of alpha or beta and gamma chains are
intermediate affinity receptors.
FUNCTION
#description
#pathway
interleukin-2 stimulated growth and differentiation of T
cells, B cells, NK cells, LAK cells, monocytes,
macrophages, and oligodendrocytes
#superfamily interleukin-2 receptor gamma chain
#cytokine_receptor; duplication; glycoprotein; transmembrane
protein
FEATURE
1-22 #domain signal sequence #status predicted #label SIG\
23-369 #product interleukin-2 receptor gamma chain #status
predicted #label MAN\
256-284 #domain transmembrane #status predicted #label TM\
71,75,84,96,159, #binding_site carbohydrate (Asn) (covalent) #status
predicted
164,306 #length 369 #molecular_weight 42241 #checksum 6734
SUMMARY
Query Match 7.2%; Score 188; DB 2; Length 369;
Best Local Similarity 26.8%; Pred. No. 1.59e-13;
Matches 51; Conservative 52; Mismatches 75; Indels 12; Gaps 11;
Db 59 EVOCVENIEYVNWSSSEPAQNLTHYRYKYSDNNFOECSHYLFSEKITSQCQIQ 118
QY 31 QIOIIEVLETVQVWNASKYSR-TNLTFHYRFN-GDE-AYDQCTNYLLQEGHTSGCLLD 87
Db 119 KEDILQYOTFVQLODPQKQPRQVQKLNQNLQVPRAPENLTLSNLSOLELRWKS 178
QY 88 AQRODILYFSIR-NGTH-PVFTASRWV-VYVL-KPSSPKHVRFS-WHODAVTV--TCS 140
Db 179 IKERQYLYQVRSNDRSWEITLVNHEPRLSPVDELKRYTFVRVRSY-NPIGSSQ 237
QY 141 LSYGDLLEYQVRSFPTDTEW-QSKQENTCNVTIEGLDAEKCYSFVVRKAMEDVYGP 199
Db 238 WSKSQPVHW 247
QY 200 PSDWSEVTCW 209
```

```
Query Match 6.8%; Score 178; DB 2; Length 467;
Best Local Similarity 30.0%; Pred. No. 7.11e-12;
Matches 24; Conservative 24; Mismatches 29; Indels 3; Gaps 3;
Db 272 ILVAVPIFLLLTGLIH-FRLSPKVKRIYQNVPSPEAFHFLYSYVHGDFQTWIGARR 330
QY 233 ILISSLAILLMSVLLSLSLKWL-RVKKFLIPSPDPKSIFFGLFEIHQGNFQEWITDQ 291
Db 331 AGPOAQ-DGASAPSGDS 349
QY 292 NVAILHKMAGAEQESGPEEP,311
RESULT 5
ENTRY
TITLE
ORGANISM
DATE
#type complete
#gene interleukin-9 receptor precursor - human
#formal_name Homo sapiens #common_name man
#sequence_revision 27-Jun-1994 #text_change
10-Sep-1997
ACCESSIONS
REFERENCE
#authors
#journal
#title
#cross-references MUID:92302307
#status preliminary
#molecule_type mRNA
#residues 1-522 #label REN
#cross-references GB:M84747; NID:gl84508; PID:gl84509
KEYWORDS
glycoprotein; receptor; T-cell proliferation; transmembrane
protein
SUMMARY
#length 522 #molecular_weight 57333 #checksum 2693
Query Match 6.8%; Score 176; DB 2; Length 522;
Best Local Similarity 27.6%; Pred. No. 1.51e-11;
Matches 30; Conservative 43; Mismatches 74; Indels 14; Gaps 12;
Db 224 LRQVATLEDDVVEERYTGQNSE--NSOPVCFQAFQPGQLIPPGWPGN-TLVAVSI 279
QY 183 FVVRKAMED-VGPTTFSDMSEVTCWQGE-IRDACAETP-TPPKKLSKFLISSLA 239
Db 280 FLILGTPTYL-LFKSPRVKRIYQNVPSPEAFHFLYSYVHGDFQTWIGARRAGVLLSQ 338
QY 240 ILLMSVLLSLSLKWL-RVKKFLIPSPDPKSIFFGLFEIHQGNFQEWITDQNVAIL-H 297
Db 339 DCAGTPQGALEPCVQATALLTCGPAPKMSVALEEEQEGPG-T-RIPGN-LSEEDVLP 395
QY 298 KMAGAEQES-GPEPLVVLQAKTEASPRMLDPQTEKEASGSLQLOPHOPLOGGDVVTI 356
Db 396 G 396
QY 357 G 357
RESULT 6
ENTRY
TITLE
ORGANISM
DATE
#type complete
#gene interleukin-9 receptor precursor - mouse
#formal_name Mus musculus #common_name house mouse
#sequence_revision 27-Jun-1994 #text_change
10-Sep-1997
ACCESSIONS
REFERENCE
#authors
#journal
#title
#cross-references MUID:92302307
#status preliminary
#molecule_type mRNA
#residues 1-522 #label REN
#cross-references GB:M84747; NID:gl84508; PID:gl84509
KEYWORDS
glycoprotein; receptor; T-cell proliferation; transmembrane
protein
SUMMARY
#length 522 #molecular_weight 57333 #checksum 2693
Query Match 6.8%; Score 176; DB 2; Length 522;
Best Local Similarity 27.6%; Pred. No. 1.51e-11;
Matches 30; Conservative 43; Mismatches 74; Indels 14; Gaps 12;
Db 224 LRQVATLEDDVVEERYTGQNSE--NSOPVCFQAFQPGQLIPPGWPGN-TLVAVSI 279
QY 183 FVVRKAMED-VGPTTFSDMSEVTCWQGE-IRDACAETP-TPPKKLSKFLISSLA 239
Db 280 FLILGTPTYL-LFKSPRVKRIYQNVPSPEAFHFLYSYVHGDFQTWIGARRAGVLLSQ 338
QY 240 ILLMSVLLSLSLKWL-RVKKFLIPSPDPKSIFFGLFEIHQGNFQEWITDQNVAIL-H 297
Db 339 DCAGTPQGALEPCVQATALLTCGPAPKMSVALEEEQEGPG-T-RIPGN-LSEEDVLP 395
QY 298 KMAGAEQES-GPEPLVVLQAKTEASPRMLDPQTEKEASGSLQLOPHOPLOGGDVVTI 356
Db 396 G 396
QY 357 G 357
```

```

##status      preliminary
##molecule_type mRNA
##residues    1-468 ##label REN
##cross-references GB:M84746; NID:g194044; PID:g194045
##note        authors translated the codon GGG for residue 394 as Glu
              glycoprotein; receptor; T-cell proliferation; transmembrane
              protein
KEYWORDS      #length 468 #molecular-weight 52260 #checksum 48
SUMMARY

Query Match      6.4%; Score 166; DB 2; Length 468;
Best Local Similarity 24.7%; Pred. No. 6.18e-10;
Matches 20; Conservative 28; Mismatches 31; Indels 2; Gaps 2;
Db 273 ILVVPIFLLTGFVHL-LFKLSRLKRFYQNPSPFAFPHLYSVYHGFQSWTARR 331
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 233 ILISSLAILLMVLSLLSLWKLW-RVKFKLIPSPDPKSIFFGLFEIHQGNFQEWITDTQ 291
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 332 AGPOARONGVSTSGASSESI 352
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 292 NVAHLHRMAGAEQSGPEPL 312
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

```

```

RESULT 7
ENTRY   A46713
TITLE   erythropoietin receptor precursor - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE     10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change
ACCESSIONS A46713
REFERENCE  #authors Masuda, S.; Nagao, M.; Takahata, K.; Konishi, Y.; Gallyas
           #journal J. Biol. Chem. (1993) 268:11208-11216
           #title Functional erythropoietin receptor of the cells with neural
           #cross-references M84746; NID:g194044; PID:g194045
           #accession A46713
           ##status preliminary
           ##molecule_type mRNA
           ##residues 1-507 ##label MAS
           ##cross-references GB:D13566; NID:g286209; PIDN:BAA02761.1;
           ##experimental_source PC12 and erythroid cells
           ##note sequence extracted from NCBI backbone (NCBIN:132811,
           NCBI:P132813)
CLASSIFICATION #superfamily erythropoietin receptor; cytokine receptor
KEYWORDS      cytokine receptor; glycoprotein; transmembrane protein
FEATURE
1-24          #domain signal sequence #status predicted #label SIG\
25-507        #product erythropoietin receptor #status predicted
              #label MAT\
25-249        #domain extracellular #status predicted #label EXT\
52-238        #domain cytokine receptor homology #label CRS\
250-271        #domain transmembrane #status predicted #label TM\
272-507        #domain intracellular #status predicted #label INT\
75            #binding_site carbohydrate (Asn) (covalent) #status
              predicted
SUMMARY      #length 507 #molecular-weight 55499 #checksum 7999
Query Match      6.3%; Score 164; DB 1; Length 507;
Best Local Similarity 45.2%; Pred. No. 1.29e-09;
Matches 28; Conservative 11; Mismatches 19; Indels 4; Gaps 3;
Db 247 LDPILTLTSL-ILVLISLLTLVALLSHRRTLQOKIWPGPSPESEGLFTTHKGNFOL 305
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 229 LSKFILLISSLAILLMVLSLL--LSLWKLWRVKFKFLI-PSVDPKSIFFGLFEIHQGNFQ 285
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 306 WL 307
   ||:
QY 286 WI 287
   ||:

```

```

RESULT 8
ENTRY   A32385
TITLE   erythropoietin receptor precursor, membrane-bound form -
        mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE     28-Sep-1990 #sequence_revision 05-Apr-1995 #text_change
ACCESSIONS A41686; A32385; S13249
REFERENCE  #authors Hino, M.; Tojo, A.; Misawa, Y.; Morii, H.; Takaku, F.;
           #journal Mol. Cell. Biol. (1991) 11:5527-5533
           #title Unregulated expression of the erythropoietin receptor gene
           #cross-references M84746; NID:g237036; PIDN:AAB20029.1; PID:g237037
           #accession A41686
           ##molecule_type mRNA
           ##residues 1-507 ##label HIN
           ##cross-references GB:S59388; NID:g237036; PIDN:AAB20029.1; PID:g237037
           ##experimental_source murine erythroleukemia (MEL) cell line F5-5
REFERENCE  #authors D'Andrea, A.D.; Lodish, H.F.; Wong, G.G.
           #journal Cell (1989) 57:277-285
           #title Expression cloning of the murine erythropoietin receptor.
           #cross-references M84746; NID:g237036; PIDN:AAB20029.1; PID:g237037
           #accession A32385
           ##molecule_type mRNA
           ##residues 1-507 ##label DAA
           ##cross-references GB:J04843; NID:g193090; PIDN:AAA37571.1; PID:g309219
           ##experimental_source murine erythroleukemia (MEL) cells, subclone 745
REFERENCE  #authors Kuramochi, S.; Ikawa, Y.; Todokoro, K.
           #journal J. Mol. Biol. (1990) 216:567-575
           #title Characterization of murine erythropoietin receptor genes.
           #cross-references M84746; NID:g237036; PIDN:AAB20029.1; PID:g237037
           #accession S13249
           ##molecule_type DNA; mRNA
           ##residues 1-507 ##label KUR
           ##cross-references EMBL:X53081; NID:g50861; PIDN:CAA37248.1; PID:g50862
           ##experimental_source murine erythroleukemia K-1 cells
GENETICS      39/1; 83/2; 142/1; 194/3; 246/1; 275/2; 304/3
INTRONS       #superfamily erythropoietin receptor; cytokine receptor
CLASSIFICATION #alternative splicing; cytokine receptor; glycoprotein;
              transmembrane protein
KEYWORDS
FEATURE
1-24          #domain signal sequence #status predicted #label SIG\
25-507        #product erythropoietin receptor #status predicted
              #label MAT\
25-249        #domain extracellular #status predicted #label EXT\
52-238        #domain cytokine receptor homology #label CRS\
250-271        #domain transmembrane #status predicted #label TM\
272-507        #domain intracellular #status predicted #label INT\
75            #disulfide_bonds #status predicted
              #binding_site carbohydrate (Asn) (covalent) #status
              predicted
SUMMARY      #length 507 #molecular-weight 55194 #checksum 9070
Query Match      6.3%; Score 163; DB 1; Length 507;
Best Local Similarity 46.8%; Pred. No. 1.86e-09;
Matches 29; Conservative 9; Mismatches 20; Indels 4; Gaps 3;
Db 247 LDPILTLTSL-ILVLISLLTLVALLSHRRTLQOKIWPGPSPESEGLFTTHKGNFOL 305
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 229 LSKFILLISSLAILLMVLSLL--LSLWKLWRVKFKFLI-PSVDPKSIFFGLFEIHQGNFQ 285
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 306 WL 307
   ||:
QY 286 WI 287
   ||:

```

```

RESULT          9
ENTRY           1
TITLE           1
ORGANISM        1
DATE            1

ZUHUR           #type complete
erythropoietin receptor precursor - human
#formal_name Homo sapiens #common_name man
#accession A43799
#cross-references GB:M60459; NID:gl82244; PIDN:AAAS2403.1; PID:gl82245
#journal Blood (1990) 76:24-30
#title The gene for the human erythropoietin receptor: analysis of
the coding sequence and assignment to chromosome 19p.
#cross-references MUID:90304334
#accession A60160
#status not compared with conceptual translation
#molecule_type mRNA; DNA
#residues 1-101; 'R', 103-188, 'RP', 191-243, 'E', 245-508 #label WIN
#authors Noguichi, C.T.; Bae, K.S.; Chin, K.; Wada, Y.; Schechter,
A.N.; Hankins, W.D.
#journal Blood (1991) 78:2548-2556
#title Cloning of the human erythropoietin receptor gene.
#cross-references MUID:92399733
#accession A49824
#molecule_type DNA
#residues 1-508 #label NOG
#cross-references GB:A45332; NID:g255496; PIDN:AAB23271.1; PID:g255497
#experimental_source Placenta
#note sequence extracted from NCBI backbone (NCBIN:113293,
NCBIP:113294)

REFERENCE
#authors Ehrenman, K.; St. John, T.
#journal Exp. Hematol. (1991) 19:973-977
#title The erythropoietin receptor gene: cloning and identification
of multiple transcripts in an erythroid cell line OCIM1.
#cross-references MUID:91372359
#accession A53958
#molecule_type mRNA
#residues 1-508 #label EHR
#authors Penny, L.A.; Forget, B.G.
#journal Genomics (1991) 11:974-980
#title Genomic organization of the human erythropoietin receptor
gene.
#cross-references MUID:92147143
#accession A55280
#molecule_type DNA
#residues 1-17; 381-387, 'LLEQOQDA', 391-395; 504-508 #label PEN
#note sequence modified after extraction from NCBI backbone
#note the authors translated the codon GAT for residue 31 as E
#note an insert compared to other published sequences is
considered by authors as likely cloning artifact
rather than polymorphism

REFERENCE
#authors Maouche, L.; Tournamille, C.; Hattab, C.; Boffa, G.; Carttron,
J.P.; Chretien, S.
#journal Blood (1991) 78:2557-2563
#title Cloning of the gene encoding the human erythropoietin
receptor.
#cross-references MUID:92399734
#accession I52563

#status translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-96 #label RES
#cross-references GB:M76595; NID:gl82147; PIDN:AAAS2393.1; PID:g553281
GENETICS
#gene GDB:EPOR
#cross-references GDB:125242; OMIM:133171
#map_position 19p13.3-19p13.2
#introns 39/1; 84/2; 143/1; 195/3; 247/1; 276/2; 305/3
CLASSIFICATION
#superfamily erythropoietin receptor; cytokine receptor
homology
KEYWORDS
alternative splicing; cytokine receptor; glycoprotein;
transmembrane protein
FEATURE
1-24 #domain signal sequence #status predicted #label SIG\
25-508 #product erythropoietin receptor #status predicted
#label MAT\
52-250 #domain extracellular #status predicted #label EXT\
52-239 #domain cytokine receptor homology #label CRS\
233-237 #region WSXWS motif\
251-272 #domain transmembrane #status predicted #label TMM\
273-508 #domain intracellular #status predicted #label INT\
#disulfide_bonds #status predicted\
52-62,91-107 #binding_site carbohydrate (Asn) (covalent) #status
predicted\
SUMMARY #length 508 #molecular_weight 55065 #checksum 8723
Query Match 6.2%; Score 160; DB 1; Length 508;
Best Local Similarity 42.6%; Pred. No. 5.53e-09;
Matches 26; Conservative 12; Mismatches 21; Indels 2; Gaps 2;
DB 248 LDPLITLSLIDVILVLLTVALSHRRALKOKIWGPSPSEFEGLFTTHKGNFQW 307
QY 229 LSKFIISSLAIIIMVSLI-LLSLWKLWRVKKFLI-PSVPDPKSIFFGLFEIHQGNFQW 286
DB 308 L 308.
QY 287 I 287
RESULT 10
ENTRY #type complete
TITLE Interleukin-3 receptor beta chain precursor - mouse
ORGANISM Mus musculus #common_name house mouse
DATE 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change
22-Jun-1999
ACCESSIONS A40091; A43022
REFERENCE A40091
#authors Itoh, N.; Yonehara, S.; Schreurs, J.; Gorman, D.M.; Maruyama,
K.; Ishii, A.; Yahara, I.; Arai, K.; Miyajima, A.
#journal Science (1990) 247:324-327
#title Cloning of an interleukin-3 receptor gene: a member of a
distinct receptor gene family.
#cross-references MUID:90117145
#accession A40091
#status nucleic acid sequence not shown
#residues 1-878 #label ITO
REFERENCE A43022
#authors Gorman, D.M.
#submission submitted to GenBank, November 1989
#accession A43022
#molecule_type mRNA
#residues 1-815; 'Q', 817-878 #label GOR
#cross-references GB:M29855; NID:gl98342; PIDN:AAAS39295.1; PID:g309406
COMMENT In mice there are two classes of high-affinity IL-3 receptors. One
contains this IL-3-specific beta chain and the other contains the
beta chain also shared by high-affinity IL-5 and GM-CSF
receptors.
CLASSIFICATION #superfamily interleukin-3 receptor beta chain; cytokine
receptor homology
KEYWORDS cytokine receptor; duplication; transmembrane protein
FEATURE

```

QY	206	VTQWGEI	214
RESULT	12		
ENTRY			
TITLE			
ORGANISM			
DATE			
ACCESSIONS			
REFERENCE			
#authors			
#journal			
#title			
#cross-references			
#accession			
#status			
#molecule_type			
#residues			
#cross-references			
#experimental_source			
GENETICS			
#gene			
SUMMARY			
Query Match			
Best Local Similarity			
Matches			
Db			
QY			
RESULT	13		
ENTRY			
TITLE			
ORGANISM			
DATE			
ACCESSIONS			
REFERENCE			
#authors			
#journal			
#title			
#cross-references			
#accession			
#status			
#molecule_type			
#residues			
#cross-references			
#experimental_source			
GENETICS			
#gene			
SUMMARY			
Query Match			

```

Best Local Similarity 35.1%; Pred. NO. 1.41e-01;
Matches 20; Conservative 11; Mismatches 21; Indels 5; Gaps 5;

Db 206 AVLVICDD-AYSGFEYEAASLRMSGFARFAQAHNKNCALIKDLTKEE-YA-NGLRV 259
||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
133 AVTVTCSDLSYGDLLYEQY-RSPDFEWTQSQKQENTCVTIEGLDAKCYSFV-VRV 187

RESULT 14
ENTRY B28516 #type complete
TITLE cytochrome P450 2C7 - rat
ALTERNATE_NAMES cytochrome P450 1F-1; cytochrome P450f; P450-RLM5b
CONTAINS oxidoreductase (EC 1.-.-.-)
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 15-Dec-1988 #sequence_revision 31-Dec-1993 #text_change
11-Jun-1999
ACCESSIONS B28516; B25954; A60563; S24582; A25585; A27425
REFERENCE B28516
AUTHORS Kimura, H.; Yoshioka, H.; Sogawa, K.; Sakai, Y.;
Fujii-Kuriyama, Y.
J Biol Chem (1988) 263:701-707
#journal Complementary DNA cloning of cytochrome P-450s related to
#title P-450(M-1) from the complementary DNA library of female rat
livers. Predicted primary structures for P-450f, PB-1, and
PB-1-related protein with a bizarre replacement block and
their mode of transcriptional expression.
#cross-references MUID:88087187
#accession B28516
#molecule_type mRNA
#residues 1-102, N', 104-445 #label KIM
#cross-references EMBL:M18774
REFERENCE A92552
AUTHORS Gonzalez, F.J.; Kimura, S.; Song, B.J.; Pastewka, J.;
Gelboin, H.V.; Hardwick, J.P.
J. Biol. Chem. (1986) 261:10667-10672
#journal Sequence of two related P-450 mRNAs transcriptionally
#title increased during rat development. Sequence occupies the
complete 3' untranslated region of a liver mRNA.
#cross-references MUID:86278140
#accession B25954
#molecule_type mRNA
#residues 1-278, 'H', 280-350, 'HD', 353-490 #label GON
#note the authors translated the codon CAT for residue 279 as
Gln, CTG for residue 313 as Met, ATG for residue 314
as Leu, CAT for residue 351 as Met, GAT for residue
352 as Ile, and GAG for residue 415 as Gly
REFERENCE A60563
AUTHORS Westin, S.; Strom, A.; Gustafsson, J.A.; Zaphiropoulos, P.G.
Mol. Pharmacol. (1990) 38:192-197
#journal Growth hormone regulation of the cytochrome P-450IIC
#title subfamily in the rat: inductive, repressive, and
transcriptional effects on P-450f (IIC7) and P-450-PB1
(IIC6) gene expression.
#cross-references MUID:90348631
#accession A60563
#molecule_type DNA
#residues 1-56 #label WES
#cross-references GB:X12595
REFERENCE S24582
AUTHORS Strom, A.
#submission submitted to the EMBL Data Library, August 1988
#accession S24582
#molecule_type DNA
#residues 1-56, 'V', 58, 'IVGD', 63-64, 'I', 66 #label STR
#cross-references EMBL:X12595; NID:957546; PIDN:CAA31108.1; PID:957547
#note the mismatched portion of this report appears to be
translated intron
REFERENCE A90511
AUTHORS Friedberg, T.; Waxman, D.J.; Atchison, M.; Kumar, A.;
Haapraanta, T.; Raphael, C.; Adesnik, M.
#journal Biochemistry (1986) 25:7975-7983
#title Isolation and characterization of cDNA clones for cytochrome
P-450 immunochemically related to rat hepatic P-450 form

```

```

PB-1.
#cross-references MUID:871010195
#accesion A25585
#molecule_type mRNA
#residues 412-470, 'A', 472-490 ##label PRI
REFERENCE
A92615 Pavreau, J.V.; Malchoff, D.M.; Mole, J.E.; Schenkman, J.B.
J. Biol. Chem. (1987) 262:14319-14326
Responses to insulin by two forms of rat hepatic microsomal
cytochrome P-450 that undergo major (RLM6) and minor
(RLMSD) elevations in diabetes.
#cross-references MUID:88007689
#accesion A27425
#molecule_type protein
#residues 1-23 ##label FAV
GENETICS
CYP2C7
CLASSIFICATION
superfamily human cytochrome P450 CYP2B6; cytochrome P450
homology
KEYWORDS
chromoprotein; electron transfer; endoplasmic reticulum;
heme; iron; monooxygenase; oxidoreductase; transmembrane
protein
FEATURE
435 #binding_site heme iron (Cys) (axial ligand) #status
predicted
SUMMARY
length 490 #molecular-weight 56186 #checksum 8804
Query Match 4.2%; Score 108; DB 2; Length 490;
Best Local Similarity 32.7%; Pred. No. 1.90e+01;
Matches 18; Conservative 17; Mismatches 18; Indels 2; Gaps 2;
Db 3 LVTF-LVLTSLLSLLWSQSSRRKTPCGTPTPLIIGNFLQIDVKNIQSFLT 55
|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 234 LISSALLLVSVLLSLLWKLVRKKFIPSPDPKSIIFPGLFEIHQNFOEWIT 288
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
RESULT 15
ENTRY
TITLE Interleukin-2 receptor beta chain precursor - human
ALTERNATE_NAMES CD25 beta chain
ORGANISM Homo sapiens
COMMON_NAME man
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
20-Mar-1998
ACCESSIONS
A30342
REFERENCE
Hatakeyama, M.; Tsudo, M.; Minamoto, S.; Kono, T.; Doi, T.;
Miyata, T.; Miyasaka, M.; Taniguchi, T.
Science (1989) 244:551-556
Interleukin-2 receptor beta-chain gene: generation of three
receptor forms by cloned human alpha- and beta-chain
cDNA's.
#cross-references MUID:89242117
#accesion A30342
#molecule_type mRNA
#residues 1-551 ##label HAT
#cross-references GB:M26062; NID:g186322; PID:g307048
GENETICS
GDB:IL2RB
#gene IL2RB
#cross-references GDB:I18822; OMIM:146710
map_position 22q11.2-22q13
KEYWORDS
cytokine receptor; transmembrane protein
FEATURE
1-26 : domain signal sequence #status predicted #label SIG\
27-551 : product interleukin-2 receptor beta chain #status
predicted #label MAR
SUMMARY
length 551 #molecular-weight 6117 #checksum 5946
Query Match 4.2%; Score 110; DB 2; Length 551;
Best Local Similarity 26.7%; Pred. No. 1.05e+01;
Matches 32; Conservative 28; Mismatches 52; Indels 8; Gaps 8;
Db 193 ICLETLTPDTQEYEQVVRKPLOGEEFT-TWSP-WSQPLAF-RPK-PAAIGKDPIPWIGHL 247

```

QY 170 VTIEGLDAEKCYSEWVRVKAMEDYIGPTYPDSEVTCWORGEIRDACAETPTPPKPL 229  
Db 248 L-VGLSGAFGEIILVYLLINCRNTGPW-LKKVLKCNTPDPKSKFFSOLSSEHGDDVQKWLS 305  
QY 230 SKFILISSLAILLMVSLLLSLWK-LWRVKRFLIPSVDPDKSIFPGLFEIHQGNFQEWIT 288

Search completed: Wed May 10 11:17:42 2000  
Job time : 28 secs.

\*\*\*\*\*

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 11:11:47 2000; MasPar time 80.57 Seconds

Tabular output not generated.

Title: >US-09-376-430-2  
Description: (23-371) from US09376430A.pap (2 of 25)  
Perfect Score: 2601  
Sequence: 1 QGGAEGVQIILYENLEIV.....DVVTGGFTFVNDRIYVAL 349

Scoring table: PAM 150  
Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot38  
1:swissprot

Statistics: Mean 48.690; Variance 85.970; scale 0.566

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description	Pred. No.
1	219	8.4	1	CYRG_BOVIN	2 34e-21
2	210	8.1	1	CYRG_CANFA	1 27e-19
3	195	7.5	1	CYRG_HUMAN	8 87e-17
4	188	7.2	1	CYRG_MOUSE	1 79e-15
5	176	6.8	1	IL2B_MOUSE	2 84e-13
6	166	6.4	1	IL2B_MOUSE	1 76e-11
7	164	6.3	1	EPOR_MOUSE	3 98e-11
8	163	6.3	1	EPOR_MOUSE	2 01e-10
9	160	6.2	1	CP23_MOUSE	9 08e-05
10	156	6.2	1	CP23_MOUSE	1 08e-03
11	154	6.2	1	IL13_MOUSE	4 30e-03
12	153	6.2	1	IL13_MOUSE	8 46e-03
13	153	6.2	1	CP24_MOUSE	8 46e-03
14	153	6.2	1	CP24_MOUSE	2 30e-02
15	153	6.2	1	CP24_MOUSE	2 30e-02
16	153	6.2	1	CP24_MOUSE	2 30e-02
17	153	6.2	1	CP24_MOUSE	2 30e-02
18	153	6.2	1	CP24_MOUSE	2 30e-02
19	153	6.2	1	CP24_MOUSE	2 30e-02
20	153	6.2	1	CP24_MOUSE	2 30e-02
21	153	6.2	1	CP24_MOUSE	2 30e-02
22	153	6.2	1	CP24_MOUSE	2 30e-02
23	153	6.2	1	CP24_MOUSE	2 30e-02

RESULT	ID	CYRG_BOVIN	STANDARD	PRT	379 AA
AC	Q95118				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	15-JUL-1998	(Rel. 36, Last annotation update)			
DE	CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C)				
DE	(INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64)				
GN	IL2RG				
OS	Bos taurus (Bovine)				
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
CC	Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;				
CC	Bovinae; Bos				
CC	[1]				
RP	SEQUENCE FROM N.A.				
RA	MEDLINE; 96268473				
RA	Yoo J., Stone R.T., Solinas-Toldo S., Fries R., Beattie C.W.;				
RT	"Cloning and chromosomal mapping of bovine interleukin-2 receptor gamma gene."				
RT	DNA Cell Biol. 15:453-459(1996).				
CC	FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF INTERLEUKINS.				
CC	PROBABLY ALSO THE IL-13 RECEPTOR.				
CC	LOCUS: SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.				
CC	SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.				
CC	SIMILARITY: CONTAINS 1 FIBROBLAST TYPE III-LIKE DOMAIN.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
DR	EMBL; U33748; AAB07812.1; -				
DR	HSSP; P31785; ILIN				
DR	PROSITE; PS00241; RECEPTOR_CYTOKINES_1; 1.				
DR	PROSITE; PS00340; RECEPTOR_CYTOKINES_2; FALSE_NEG.				
DR	PFAM; PF00041; fn3; 1.				
KW	Receptor; Transmembrane; Glycoprotein; Signal.				
FT	SIGNAL	1	22	POTENTIAL.	
FT	DOMAIN	23	379	CYTOKINE RECEPTOR COMMON GAMMA CHAIN.	
FT	TRANSMEM	23	269	EXTRACELLULAR (POTENTIAL).	
FT	DOMAIN	270	290	POTENTIAL.	
FT	DOMAIN	291	379	CYTOPLASMIC (POTENTIAL).	

ALIGNMENTS

RESULT	ID	CYRG_BOVIN	STANDARD	PRT	379 AA
AC	Q95118				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	15-JUL-1998	(Rel. 36, Last annotation update)			
DE	CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C)				
DE	(INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64)				
GN	IL2RG				
OS	Bos taurus (Bovine)				
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
CC	Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;				
CC	Bovinae; Bos				
CC	[1]				
RP	SEQUENCE FROM N.A.				
RA	MEDLINE; 96268473				
RA	Yoo J., Stone R.T., Solinas-Toldo S., Fries R., Beattie C.W.;				
RT	"Cloning and chromosomal mapping of bovine interleukin-2 receptor gamma gene."				
RT	DNA Cell Biol. 15:453-459(1996).				
CC	FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF INTERLEUKINS.				
CC	PROBABLY ALSO THE IL-13 RECEPTOR.				
CC	LOCUS: SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.				
CC	SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.				
CC	SIMILARITY: CONTAINS 1 FIBROBLAST TYPE III-LIKE DOMAIN.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
DR	EMBL; U33748; AAB07812.1; -				
DR	HSSP; P31785; ILIN				
DR	PROSITE; PS00241; RECEPTOR_CYTOKINES_1; 1.				
DR	PROSITE; PS00340; RECEPTOR_CYTOKINES_2; FALSE_NEG.				
DR	PFAM; PF00041; fn3; 1.				
KW	Receptor; Transmembrane; Glycoprotein; Signal.				
FT	SIGNAL	1	22	POTENTIAL.	
FT	DOMAIN	23	379	CYTOKINE RECEPTOR COMMON GAMMA CHAIN.	
FT	TRANSMEM	23	269	EXTRACELLULAR (POTENTIAL).	
FT	DOMAIN	270	290	POTENTIAL.	
FT	DOMAIN	291	379	CYTOPLASMIC (POTENTIAL).	

FT	DOMAIN	158	256	FIBRONECTIN TYPE-III.	
FT	DISULFID	68	78	POTENTIAL.	
FT	DISULFID	109	122	POTENTIAL.	
FT	CARBOHYD	77	77	POTENTIAL.	
FT	CARBOHYD	81	81	POTENTIAL.	
FT	CARBOHYD	90	90	POTENTIAL.	
FT	CARBOHYD	166	166	POTENTIAL.	
FT	CARBOHYD	171	171	POTENTIAL.	
SO	SEQUENCE	379 AA; 43037 MW; 33CFAD9C9B032178 CRC64;			

Query Match		8.4%;	Score 219; DB 1; Length 379;
Best Local Similarity		27.5%;	Pred. No. 2.34e-21;
Matches		74; Conservative	66; Mismatches 104; Indels 25; Gaps 23;

Db	66	VQCFVFNVEYANCNTWSSSPQNNLTHGYGNFGDDKLOPCGHYLFSEGITSGCF- 124
Qy	32	IQIIFNLETVQVTWNASKYSRTN-LTFHF--R-FNGDEAYDOCTNLLQEGHTSGCLLD 87
Db	125	G-KKEIRLYETFFVQLQDPREHRKQPKOMLKLODLVWPAPENLTLRNLSQELSLWSN 183
Qy	88	AEQRDDILY--FSIR-NGTHPVTAASRWV-YY-L-KPSPKHVRESWQD-AVTVTCSD 140
Db	184	-RYLDHCLHLVQYRSDRRDSRWTEQQSVDRHHSFLSPVDAQKLYTFVRSRY-NPLCGSA 241
Qy	141	LSVGD-L-YEVOYRSPFDTEW-QSKQENTCNVTIEGLDAEKCYSEWVRKAMEDVYGD 197
Db	242	QHWSDSYPIHWGNSNTSKENIENPENPSLFALEA-VLIP-LGSMGLI-VSLTCTVC-W-L 296
Qy	198	TYPSDSEVTCWGEIRDCATPTPPKPKLSKFILISLAILMVSLLLLLSLKLNRV 257
Db	297	ERTW-PRIPTLKNL-EDLVTEYQGNFSAW 323
Qy	258	KKEILPSVDPKGIFFGLFPIHONGFOW 286

```

RESULT      2
ID          CYRG_CANFA      STANDARD;      PRT;      373 AA.
AC          PA0321;
DT          01-FEB-1995 (Rel. 31, Created)
DT          01-FEB-1995 (Rel. 31, Last sequence update)
DT          13-JUL-1998 (Rel. 36, Last annotation update)
DE          CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C)
DE          (INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64).
DE          IL2RG.
OS          Canis familiaris (Dog).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC          Eutheria; Carnivora; Fissipedia; Canidae; Canis.
[1]
RN          SEQUENCE FROM N.A.
RP          TISSUE=SPLEEN;
RX          MEDLINE; 95130114.
RX          Henthorn P.S.; Somborg R.L., Finiani V.M., Puck J.M., Patterson D.F.,
RA          Felsburg P.J.;
RT          "IL-2R gamma gene microdeletion demonstrates that canine X-linked
RT          severe combined immunodeficiency is a homologue of the human
RT          disease."
RT          Genomics 23:69-74(1994).
CC          -!- FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF
CC          INTERLEUKINS.
CC          -!- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND
CC          PROBABLY ALSO THE IL-13 RECEPTORS.
CC          -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC          -!- DISEASE: DEFECTS IN IL2RG ARE THE CAUSE OF A CANINE X-LINKED
CC          SEVERE COMBINED IMMUNODEFICIENCY.
CC          -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
CC          -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC          -----
CC          This SWISS-PROT entry is copyright. It is produced through a collaboration
CC          between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC          the European Bioinformatics Institute. There are no restrictions on its
CC          use by non-profit institutions as long as its content is in no way
CC          modified and this statement is not removed. Usage by and for commercial
CC          entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U04361; AAC48403.1; -.
DR HSSP; P31785; 1ILN.
DR PROSITE; PS00241; RECEPTOR_CYTOKINES_1; 1.
DR PROSITE; PS00340; RECEPTOR_CYTOKINES_2; FALSE_NEG.
DR PFAM; PF00041; fn3; 1.
KW Receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 373 CYTOKINE RECEPTOR COMMON GAMMA CHAIN.
FT DOMAIN 23 261 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 262 283 POTENTIAL.
FT DOMAIN 284 373 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 151 249 FIBRONECTIN TYPE-III.
FT DISULFID 62 72 POTENTIAL.
FT DISULFID 102 115 POTENTIAL.
FT CARBOHYD 24 24 POTENTIAL.
FT CARBOHYD 71 71 POTENTIAL.
FT CARBOHYD 75 75 POTENTIAL.
FT CARBOHYD 84 84 POTENTIAL.
FT CARBOHYD 159 159 POTENTIAL.
FT CARBOHYD 164 164 POTENTIAL.
FT CARBOHYD 249 249 POTENTIAL.
SQ SEQUENCE 373 AA; 42516 MW; 03A0DE1FB08D9B8 CRC64;

Query Match      8.1% Score 210; DB 1; Length 373;
Best Local Similarity 25.9%; Pred.No. 1.27e-19;
Matches 49% Conservative 56; Mismatches 73; Indels 11; Gaps 1;

Db 59 EVQCFVFNVEYMCNTNSSEPRPTNLTHYYKNSNDKVOECGHYLFPSREVTAGCWQLQ 118
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 31 QIQIIFYENLETQVWNASKYSR-TNLTFHRYE-NGDEA-YDOCTNYLLQEGHTSGCLLD 87

Db 119 KEETHLYETFWOLRDPREPROSTOKKLQNIVTPWAPENTLNLSSESLEQSWSNRH 178
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 88 AEQRD--DILYIFSRINGTHPVFTS-RWVVYL-KPSPKHVR-FSWHQDAVTVCSDLS 142

Db 179 LDHGLEHVQYRSWDPRSQTQQSVDRHSFSLPSVDGQFYFRVSRY-NPLCGSAQRW 237
   | | | | | | | | | | | | | | | | | | | | | | | | : | : | : | : |
Qy 143 YGD-LIAEQVYRSPFTEW-QSKQENTCVNTIEGLDAEKCYFWRVKAMEDVYGPDTP 200

Db 238 SEWSHPHW 246
   | : | | |
Qy 201 SDWSEVTCW 209

RESULT 3
ID CYRG-HUMAN STANDARD; PRT; 369 AA.
AC P31785;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C) (INTERLEUKIN-
DE 2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64) (CD132 ANTIGEN).
GN IL2RG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
[1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE; 92335883.
RA Takeishi T., Asao H., Ohtani K., Ishii N., Kumaki S., Tanaka N.,
RA Munakata H., Nakamura M., Sugamura K.;
RT "Cloning of the gamma chain of the human IL-2 receptor.";
RL Science 257:379-382(1992).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=LIVER.
RX MEDLINE; 93293887.
RA Noguchi M., Adelstein S., Cao X., Leonard W.J.;
RT "Characterization of the human interleukin-2 receptor gamma chain
RT gene.";
RL J. Biol. Chem. 268:13601-13608(1993).
```



RT RP SEQUENCE FROM N.A., AND VARIANTS ASP-114 AND ASN-153.  
RX MEDLINE; 94004847.  
RA Puck J.M., Deschenes S.M., Porter J.C., Dutra A.S., Brown C.J.,  
RA Willard H., Henthorn P.S.;  
RT "The interleukin-2 receptor gamma chain maps to xq13.1 and is mutated  
RT in X-linked severe combined immunodeficiency, SCIDX1.";  
RL Hum. Mol. Genet. 2:1099-1104(1993).  
[4]  
RN IDENTIFICATION AS A IL-4R SUBUNIT.  
RX MEDLINE; 94090315.  
RA Kondo M., Takeshita T., Ishii N., Nakamura M., Watanabe S.,  
RA Arai K.-I., Sugamura K., Harada N., Nakamura Y., Noguchi M.,  
RT "Sharing of the interleukin-2 (IL-2) receptor gamma chain between  
RT receptors for IL-2 and IL-4.";  
RL Science 262:1874-1877(1993).  
[5]  
RN IDENTIFICATION AS A IL-4R SUBUNIT.  
RX MEDLINE; 94090317.  
RA Russell S.M., Kegan A.D., Harada N., Nakamura Y., Noguchi M.,  
RA Leland P., Friedmann M.C., Miyajima A., Puri R.K., Paul W.E.,  
RA Leonard W.J.;  
RT "Interleukin-2 receptor gamma chain: a functional component of the  
RT interleukin-4 receptor.";  
RL Science 262:1880-1883(1993).  
[6]  
RN IDENTIFICATION AS A IL-7R SUBUNIT.  
RX MEDLINE; 94090316.  
RA Noguchi M., Nakamura Y., Russell S.M., Ziegler S.F., Tsang M., Cao X.,  
RA Leonard W.J.;  
RT "Interleukin-2 receptor gamma chain: a functional component of the  
RT interleukin-7 receptor.";  
RL Science 262:1877-1880(1993).  
[7]  
RN 3D-STRUCTURE MODELING OF 57-248.  
RX MEDLINE; 95111955.  
RA Bamorough P., Hedgecock C.J., Richards W.G.;  
RT "The interleukin-2 and interleukin-4 receptors studied by molecular  
RT modelling.";  
RL Structure 2:839-851(1994).  
[8]  
RN VARIANTS XSCID PHE-115; CYS-240 AND ILE-241.  
RX MEDLINE; 94130970.  
RA Disanto J.P., Dautry-Varsat A., Certain S., Fischer A.,  
RA de Saint Basile G.;  
RT "Interleukin-2 (IL-2) receptor gamma chain mutations in X-linked  
RT severe combined immunodeficiency disease result in the loss of  
RT high-affinity IL-2 receptor binding.";  
RL Eur. J. Immunol. 24:475-479(1994).  
[9]  
RN VARIANT XSCID LYS-68.  
RX MEDLINE; 94375038.  
RA Markiewicz S., Subtil A., Dautry-Varsat A., Fischer A.,  
RA de Saint Basile G.;  
RT "Detection of three nonsense mutations and one missense mutation in  
RT the interleukin-2 receptor gamma chain gene in SCIDX1 that  
RT differently affect the mRNA processing.";  
RL Genomics 21:291-293(1994).  
[10]  
RN VARIANT XSCID HIS-162.  
RX MEDLINE; 94300093.  
RA Ishii N., Asao K., Kimura Y., Takeshita T., Nakamura M., Tsuchiya S.,  
RA Konno T., Maeda M., Uchiyama T., Sugamura K.;  
RT "Impairment of ligand binding and growth signaling of mutant IL-2  
RT receptor gamma-chains in patients with X-linked severe combined  
RT immunodeficiency.";  
RL J. Immunol. 153:1310-1317(1994).  
[11]  
RN VARIANT XSCID ASN-39.  
RX MEDLINE; 95023932.  
RA Disanto J.P., Rieux-Laucat F., Dautry-Varsat A., Fischer A.,  
RA de Saint Basile G.;  
RT "Defective human interleukin 2 receptor gamma chain in an atypical X  
RT

RT cells";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:9456-9470(1994).  
[12]  
RN VARIANTS XSCID CYS-226 AND HIS-226.  
RX MEDLINE; 95397841.  
RA Pepper A.E., Buckley R.H., Small T.N., Puck J.M.;  
RT "Two mutational hotspots in the interleukin-2 receptor gamma chain  
RT gene causing human X-linked severe combined immunodeficiency.";  
RL Am. J. Hum. Genet. 57:564-571(1995).  
[13]  
RN VARIANT XSCID SER-183.  
RX MEDLINE; 96013903.  
RA Clark P.A., Lester T., Genet S., Jones A.M., Hendriks R.,  
RA Levinsky R.L., Kinnon C.;  
RT "Screening for mutations causing X-linked severe combined  
RT immunodeficiency in the IL-2R gamma chain gene by single-strand  
RT conformation polymorphism analysis.";  
RL Hum. Genet. 96:427-432(1995).  
[14]  
RN VARIANT XSCID GLN-HIS-TRP INS-237.  
RX MEDLINE; 95164726.  
RA Puck J.M., Pepper A.E., Bedard P.-M., Laframboise R.;  
RT "Female germ line mosaicism as the origin of a unique IL-2 receptor  
RT gamma-chain mutation causing X-linked severe combined  
RT immunodeficiency.";  
RL J. Clin. Invest. 95:895-899(1995).  
[15]  
RN VARIANT XSCID GLN-271.  
RX MEDLINE; 95190013.  
RA Schmalstieg F.C., Leonard W.J., Noguchi M., Berg M., Rudloff H.E.,  
RA Denney R.M., Dave S.K., Brooks E.G., Goldman A.S.;  
RT "Missense mutation in exon 7 of the common gamma chain gene causes a  
RT moderate form of X-linked combined immunodeficiency.";  
RL J. Clin. Invest. 95:1169-1173(1995).  
[16]  
RN VARIANT XSCID ARG-115.  
RX MEDLINE; 97042245.  
RA Stephan V., Wain V., le Deist F., Dirksen U., Broker B.,  
RA Muller-Fleckenstein I., Horneff G., Schroten H., Fischer A.,  
RA de Saint Basile G.;  
RT "Atypical X-linked severe combined immunodeficiency due to possible  
RT spontaneous reversion of the genetic defect in T cells.";  
RL New Engl. J. Med. 335:1563-1567(1996).  
[17]  
RN VARIANT XSCID GLN-285.  
RX MEDLINE; 97295088.  
RA Jones A.M., Clark P.A., Katz F., Genet S., McMahon C., Alterman L.,  
RA Cant A., Kinnon C.;  
RT "B-cell-negative severe combined immunodeficiency associated with a  
RT common gamma chain mutation.";  
RL Hum. Genet. 99:677-680(1997).  
[18]  
RN VARIANT XSCID CYS-222.  
RX MEDLINE; 98064061.  
RA Sharfe N., Shahar M., Roifman C.M.;  
RT "An interleukin-2 receptor gamma chain mutation with normal thymus  
RT morphology.";  
RL J. Clin. Invest. 100:3036-3043(1997).  
[19]  
RN FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF  
CC INTERLEUKINS.  
CC  
CC -1- SUBUNIT: THE GAMMA-CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND  
CC PROBABLY ALSO THE IL-13 RECEPTORS.  
CC  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC  
CC -1- DISEASE: DEFECTS IN IL2RG ARE THE CAUSE OF A SEVERE COMBINED  
CC IMMUNODEFICIENCY, WHICH IS KNOWN AS AGAMAGLOBULINEMIA, SWISS TYPE  
CC OR X-LINKED SEVERE COMBINED IMMUNODEFICIENCY DISEASE (XSCID).  
CC  
CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
CC  
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
CC  
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD132 entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd132.htm".  
CC  
CC -1- DATABASE: NAME=IL2RGbase; NOTE=X-linked SCID mutation database;  
CC WWW="http://www.nhgri.nih.gov/DIR/SCID/IL2RGbase.html".  
CC







RT "Structure and transcription of the mouse erythropoietin receptor  
 RT gene";  
 RL M1. Cell. Biol. 10:3675-3682(1990).  
 RN [5]  
 RP SEQUENCE OF 1-24 FROM N.A.  
 RX MEDLINE: 91201346.  
 RA Lacombe C., Chretien S., Lemarchandel V., Mayeux P., Romeo P.H.,  
 RA Gisselbrecht S., Cartion J.P.;  
 RT "Spleen focus-forming virus long terminal repeat insertion  
 RT activation of the murine erythropoietin receptor gene in the T3C1-2  
 RT friend leukemia cell line";  
 RL J. Biol. Chem. 266:6952-6956(1991).  
 RN [6]  
 RP MUTAGENESIS.  
 RX MEDLINE: 93180826.  
 RA Miura O., Cleveland J.L., Ihle J.N.;  
 RT "Inactivation of erythropoietin receptor function by point mutations  
 RT in a region having homology with other cytokine receptors";  
 RL Mol. Cell. Biol. 13:1788-1795(1993).  
 CC -!- FUNCTION: RECEPTOR FOR ERYTHROPOIETIN. MAY PLAY A ROLE IN THE  
 CC MECHANISM OF ERYTHROPOIETIN-INDUCED ERYTHROBLAST PROLIFERATION  
 CC AND DIFFERENTIATION.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J04843; AAA37571.1; -  
 DR EMBL: X53081; CAA37248.1; -  
 DR EMBL: M38133; AAA37572.1; -  
 DR EMBL: M62360; AAA37582.1; -  
 DR EMBL: M59388; CAB31799.1; -  
 DR PIR: A32385; A32385.  
 DR PIR: A41686; A41686.  
 DR PIR: S13249; S13249.  
 DR PIR: S14081; S14081.  
 DR HSP: P19235; 1EBP.  
 DR MGD: MGI:95408; EPOR.  
 DR PROSITE: PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
 DR PROSITE: PS00340; RECEPTOR\_CYTOKINES\_2; 1.  
 DR PFAM: PF00041; fn3; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 507  
 FT DOMAIN 25 507 ERYTHROPOIETIN RECEPTOR.  
 FT TRANSMEM 25 249 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 250 272 POTENTIAL.  
 FT DOMAIN 273 507 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 147 212 FIBRONECTIN TYPE-III.  
 FT DISULFID 52 62 BY SIMILARITY.  
 FT DISULFID 90 106 BY SIMILARITY.  
 FT CARBOHYD 75 75 POTENTIAL.  
 FT CONFLICT 291 291 E -> D (IN CAB31799).  
 SQ SEQUENCE 507 AA; 55194 MW; 067657A2E26451CA CRC64;  
 Query Match 6.3%; Score 163; DB 1; Length 507;  
 Best Local Similarity 46.8%; Pred. No. 5.98e-11;  
 Matches 29; Conservative 9; Mismatches 20; Indels 4; Gaps 3;  
 DB 247 LDPPLTSLI-LIVLISLLTLVALLSHRRTLOOKTWPGIPSPESFEGLFTTHKGNFOL 305  
 QY 229 LSKFILLISLAILLMVSLLL-LLSLKLWVRVKFELI-PSVPDPKSIFFPLGFEIHQGNFQE 285  
 DB 306 WL 307  
 QY 286 WI 287

RESULT 9  
 ID EPOR HUMAN STANDARD; PRT; 508 AA.  
 AC P19235; 1990 (Rel. 16, Created)  
 DI 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 13-DEC-1998 (Rel. 37, Last annotation update)  
 DE ERYTHROPOIETIN RECEPTOR PRECURSOR (EPO-R).  
 GN EPOR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 91372359.  
 RA Ehrenman K., St John T.;  
 RT "The erythropoietin receptor gene: cloning and identification of  
 RT multiple transcripts in an erythroid cell line OCIM1";  
 RL Exp. Hematol. 19:973-977(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 90304340.  
 RA Jones S.S., D'Andrea A., Haines L.L., Wong G.G.;  
 RT "Human erythropoietin receptor: Cloning, expression, and biologic  
 RT characterization";  
 RL Blood 76:31-35(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 92399733.  
 RA Noguchi C.T., Bae K.S., Chin K., Wada Y., Schechter A.N.,  
 RA Hankins W.D.;  
 RT "Cloning of the human erythropoietin receptor gene";  
 RL Blood 78:2548-2556(1991).  
 RN [4]  
 RP SEQUENCE OF 1-96 FROM N.A.  
 RX MEDLINE: 92399734.  
 RA Maouche L., Tournamille C., Hattab C., Boffa G., Cartion J.P.,  
 RA Chretien S.;  
 RT "Cloning of the gene encoding the human erythropoietin receptor";  
 RL Blood 78:2557-2563(1991).  
 RN [5]  
 RP SEQUENCE OF 1-17 FROM N.A.  
 RX MEDLINE: 92147143.  
 RA Penny L.A., Forget B.G.;  
 RT "Genomic organization of the human erythropoietin receptor gene";  
 RL Genomics 11:974-980(1991).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 34-244.  
 RX MEDLINE: 96291992.  
 RA Livnah O., Stura E.A., Johnson D.L., Middleton S.A., Mulcahy L.S.,  
 RA Wrighton N.C., Dower W.J., Jolliffe L.K., Wilson I.A.;  
 RT "Functional mimicry of a protein hormone by a peptide agonist: the  
 RT EPO receptor complex at 2.8 A";  
 RL Science 273:464-471(1996).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 34-244.  
 RX MEDLINE: 99023198.  
 RA Livnah O., Johnson D.L., Stura E.A., Farrell F.X., Barbone F.P.,  
 RA You Y., Liu K.D., Goldsmith M.A., He W., Krause C.D., Pestka S.,  
 RA Jolliffe L.K., Wilson I.A.;  
 RT "An antagonist peptide-EPO receptor complex suggests that receptor  
 RT dimerization is not sufficient for activation";  
 RL Nat. Struct. Biol. 5:993-1004(1998).  
 CC -!- FUNCTION: RECEPTOR FOR ERYTHROPOIETIN. MAY PLAY A ROLE IN THE  
 CC MECHANISM OF ERYTHROPOIETIN-INDUCED ERYTHROBLAST PROLIFERATION  
 CC AND DIFFERENTIATION.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC AND CARCINOGENS.  
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC  
CC  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announcement/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF047725; ARDL3720.1; -  
DR MGD; MGI\_1306819; CYP2C38.  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum.  
FT BINDING 435 435 HEME (BY SIMILARITY).  
SQ SEQUENCE 490 AA; 55089 MW; C2922E7EC36A410C CRG64;  
Query Match 4.8%; Score 126; DB 1; Length 490;  
Best Local Similarity 33.9%; Pred. No. 9.08e-05;  
Matches 19; Conservative 18; Mismatches 17; Indels 2; Gaps 2;

```

Db      3  LVTEL-VLTISSLLLSLWQ-RSRGRLLPPGTPPTIcGNFLQIDVKFNQSLTN 56
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy      234 LISSLAILLMSVLLLSLWLVKRWKFLIPSDPKSIFPGLFEIHQGNFQEWITD 289

RESULT 11
ID      I131.MOUSE      STANDARD;      PRt;      424 AA.
AC      O09030;
DT      01-NOV-1997 (Rel. 35, Created)

```

DI  
DT  
DE  
DE  
DE

01-NOV-1997 (rel. 35, Last sequence update)  
15-JUL-1999 (rel. 38, Last annotation update)  
INTERLEUKIN-13 RECEPTOR ALPHA-1 CHAIN PRECURSOR (IL-13R-ALPHA-1) (IL-13RA-1)  
INTERLEUKIN-13 BINDING PROTEIN (NR4)

GN 1113RA1 OR 1113KA OR 1113K.  
OS Mus musculus (Mouse).  
CN Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

KN  
KN  
RP  
RX  
RA

[1]  
SEQUENCE FROM N.A.  
MEDLINE; 96133964.  
Hilton D.J., Zhang J.-G., Mercalf D., Alexander W.S., Nicola N.A.

RA Willison T.A.;  
RT "Cloning and characterization of a binding subunit of the interleukin  
RT 13 receptor that is also a component of the interleukin 4 receptor.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:497-501(1996).  
CC -1- FUNCTION: BINDS IL-13 WITH A LOW AFFINITY. TOGETHER WITH IL-4R-  
CC ALPHA CAN FORM A FUNCTIONAL RECEPTOR FOR IL-13. ALSO SERVES AS AN  
CC ALTERNATE ACCESSORY PROTEIN TO THE COMMON CYTOKINE RECEPTOR GAMMA  
CC CHAIN FOR IL-4 SIGNALING, BUT CANNOT REPLACE THE FUNCTION OF  
CC GAMMA-C IN ALLOWING ENHANCED IL-2 BINDING ACTIVITY (BY  
CC

-! SUBUNIT INTERLEUKIN-13 RECEPTOR IS A COMPLEX OF IL4R-ALPHA,  
IL13R-ALPHA, AND POSSIBLY OTHER COMPONENTS (BY SIMILARITY).  
CC  
CC  
CC  
CC  
CC  
CC  
CC

- TISSUE SPECIFICITY: SPLEEN, LIVER, THYMUS, HEART, LUNG, KIDNEY, TESTIS, STOMACH, BRAIN, SKIN, AND COLON; BUT NOT SKELETAL MUSCLE.
- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use.  
 CC -----

CC 4 use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

CC	EMBL; S80963; AAB50695.1; -	
DR	MGD; MGI:105052; IL13RA.	
DR	Receptor: Transmembrane	
KW	Receptor: Glycoprotein: Immunoglobulin domain: Signal	

FT SIGNAL 1 25  
FT CHAIN 26 424  
FT DOMAIN 26 340  
FT TRANSFEM 341 364  
FT DOMAIN 365 424  
FT DOMAIN 37 100  
FT DOMAIN 44 93  
FT DISULFID 132 142  
FT DISULFID 171 182  
FT DISULFID 35 35  
FT CARBOHYD 59 59  
FT CARBOHYD 103 103  
FT CARBOHYD 136 136  
FT CARBOHYD 262 262  
FT CARBOHYD 338 338  
SQ SEQUENCE 424 AA; 48402 MW; EB8330A0DC82C9F9 CRC64;

Query Match 4.6%; Score 119; DB 1; Length 424;  
Best Local Similarity 20.4%; Pred. No. 1.08e-03;  
Matches 58; Conservative 88; Mismatches 110; Indels 29; Gaps 26;

Db 120 EGDPSAVTELKCIWHNLNLSYMKCSWLPGRNTSPDTHLYTYWISLEKSRQCN-IYREG 178  
QY 23 QGGAEGV-QIQLIYFLETQVVTWNSKY-SR-TNLTFHYRFNGDEAYDOCTNLLQEG 79  
Db 179 QHIACSFKLTVKSPFHFONQVIMVKNAGKIRPCKIVSLTSYVKKDPP-HIKHLLKN 237  
QY 80 HTSGC-L-LD-AQR-D-DILYFSIRNTHPVFTASR--WVYILKPSPKHVR-FSWHQ 131  
Db 238 GALLVQKNPNFONRCITVEVNNVT-OTDRHNILEVEEDKCNQSDRNWEGTSCFQL 296  
QY 132 DAVVTTCSD-LSY-GD-LIYEVQKSPDTEWQS--K-QENFC-VNTIE-GDAEKCYSF 183  
Db 297 -PGVLA-DAVTVRY-RVK-TNKLCDNDKLSWDSWAEQSGKQNSFTYTMILTIVF 352  
QY 184 WYVKAMEDVYGPDTYPSDWSEVTCWQGEINDACAETPTPKPLSKF--ILTSILAIL 241  
Db 353 VAVAVIIL-LFYLKRLKIIIPPPDPGKIFKEMFGDQNDTLHW 396  
QY 242 LMVSLLLSLKLMWVKVFLPSVDPKSIFFGLFEIHQGNFQEW 286

RESULT 12  
ID IL3B\_MOUSE STANDARD; PRT; 878 AA.  
AC P26954;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-FEB-2000 (Rel. 39, Last annotation update)  
DE INTERLEUKIN-3 RECEPTOR CLASS II BETA CHAIN PRECURSOR (COLONY  
DE STIMULATING FACTOR 2 RECEPTOR, BETA 2 CHAIN).  
GN CSF2RB2 OR ALZCA OR IL3RB2 OR IL3R.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90117145.  
RA Itoh N., Yonehara S., Schreurs J., Gorman D.M., Maruyama K., Ishii A.,  
RA Yahara I., Arai K., Miyajima A.;  
RT "Cloning of an interleukin-3 receptor gene: a member of a distinct  
RT receptor gene family."  
RL Science 247:324-327(1990).  
CC -!- FUNCTION: IN MOUSE THERE ARE TWO CLASSES OF HIGH-AFFINITY IL-3  
CC RECEPTORS. ONE CONTAINS THIS IL-3-SPECIFIC BETA CHAIN AND THE  
CC OTHER CONTAINS THE BETA CHAIN ALSO SHARED BY HIGH-AFFINITY IL-5  
CC AND GM-CSF RECEPTORS.  
CC -!- SUBUNIT: HETEROIDIMER OF AN ALPHA AND A BETA CHAIN.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; M29855; AAA39295.1; --  
CC PIR; A40091; A40091.  
CC MGD; MGI:1339760; CSF2RB2.  
CC DR PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
CC DR PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; 1.  
CC DR PRAM; PF00041; fn3; 2.  
CC KW Receptor; Transmembrane; Glycoprotein; Signal.  
FT SIGNAL 1 22  
FT CHAIN 23 878  
FT DOMAIN 23 440  
FT TRANSFEM 441 462  
FT DOMAIN 463 878  
FT DISULFID 39 49  
FT DISULFID 78 95  
FT DISULFID 254 264  
FT DISULFID 293 310  
FT CARBOHYD 62 62  
FT CARBOHYD 350 350  
SQ SEQUENCE 878 AA; 97195 MW; 8EBC9092ADC24D56 CRC64;

Query Match 4.4%; Score 115; DB 1; Length 878;  
Best Local Similarity 32.4%; Pred. No. 4.30e-03;  
Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 370 YIDTFQYQYKKSES-WKDSKTENLGRVNSMDLPOLPDPDTSYCARVVKPSD-YD-GI 426  
QY 143 YGDLLEYQYKSPDTEW-QSRQENTCNV-TIE-G-LDREKCYSEWVRVKAMEDYVGPD 198  
Db 427 W-SENSNYTW 436  
QY 199 YPSDWSEVTCW 209

RESULT 13  
ID CP24\_MOUSE STANDARD; PRT; 490 AA.  
AC P56656;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME P450 2C39 (EC 1.14.14.1) (CYP11C39).  
GN CYP2C39.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=CD-1;  
RX MEDLINE; 96389577.  
RA Luo G., Zeidman D.C., Blaisdell J.A., Hodgson E., Goldstein J.A.;  
RT "Cloning and expression of murine CYP2C3 and their ability to  
RT metabolize arachidonic acid."  
RL Arch. Biochem. Biophys. 357:45-57(1998).  
CC -!- FUNCTION: METABOLIZES ARACHIDONIC ACID TO PRODUCE 14,15-CIS-  
CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +  
CC OXIDIZED FLAVOPROTEIN + H(2)O.  
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
CC -!- TISSUE SPECIFICITY: LIVER.  
CC -!- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER  
CC TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES,  
CC AND CARCINOGENS.  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF047726; AAD13721.1; -  
DR MGD: MGE:1306818; CYP2C39;  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum.  
FT BINDING 435 HEME (BY SIMILARITY).  
SQ SEQUENCE 490 AA; 55902 MW; 97EF56D0A4E728C3 CRC64;

Query Match 4.3%; Score 113; DB 1; Length 490;  
Best Local Similarity 33.9%;  
Pred. No. 8.46e-03;  
Matches 19; Conservative 16; Mismatches 19; Indels 2; Gaps 2;

```

Db      3  LVTF-LVTLSLLLNQSCGRSL-PPGTEFFPIIGNFLQIDMKNFSSQLTN 56
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      234  LISSAILMVLSSLLSLKWLRYVKFLIVSPDPKSIFFGLFELHQGNQOEIIT 289

```

RESULT	14	
ID	CYRB_HUMAN	PRT; 897 AA.

DT	01-OCT-1993 (Rel. 27, Created)
DT	01-FEB-1996 (Rel. 33, Last sequence update)
DT	01-NOV-1997 (Rel. 35, Last annotation update)
DE	CYTOKINE RECEPTOR COMMON BETA CHAIN PRECURSOR (CD131 ANTIGEN).
GN	CSF2RB OR IL5RB OR IL3RB.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa;  
OC Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]

RP  
SEQUENCE FROM N.A.  
RX MEDLINE; 91088571.  
RA Hayashida K., Kitamura T., Gorman D.M., Arai K., Yokota T.,  
RA Miyaajima A.;

RT "Molecular cloning of a second subunit of the receptor for human  
RT granulocyte-macrophage colony-stimulating factor (GM-CSF):  
RT reconstitution of a high-affinity GM-CSF receptor";  
RT Proc. Natl. Acad. Sci. U.S.A. 87: 9655-9659(1990)  
RL

RN [2]  
RP REVISION TO 454.  
RA Kitamura T.;  
RL Submitted (FEB-1991) to the EMBL/GenBank/DBJ databases

CC - FUNCTION: HIGH AFFINITY RECEPTOR FOR INTERLEUKIN-3, INTERLEUKIN-5  
CC AND GRANULOCYTE-MACROPHAGE COLONY-STIMULATING FACTOR.  
CC - SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. THE BETA  
CC CHAIN IS COMMON TO THE IL-3, IL-5 AND GM-CSF RECEPTORS

CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC - SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.  
CC - SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
CC - DATABASE: NAME=PROM; NOTE=CD guide CD431 structure

This SWISS-PROT entry is copyright © It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL European Molecular Biology Laboratory.

-----  
www="http://www.ncbi.nlm.nih.gov/prov/cd/cdw131.htm".

CC BY-NC-ND 4.0 International license. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

```
CC or send an email to license@slb-sib.ch).
CC -----
CC EMBL: M59941; AAA18171.1; -
CC DDB: A30255; A30358
CC DR
```

```

DR      MIN: 138981;  -
DR      PROSITE: PS00241;  RECEPTOR_CYTOKINES_1;  2.
DR      PROSITE: PS00340;  RECEPTOR_CYTOKINES_2;  1.
DR      PRM: DFG0004_5

```

Receptor; Transmembrane; Glycoprotein; Repeat; Signal.  
 SIGNAL 1 16  
 CHAIN 17 897  
 CHAIN 17 897  
 DOMAIN 17  
 CYTOKINE RECEPTOR COMMON BETA CHAIN.

1. **INTRODUCTION**

FT	TRANSMEM	444	460	POTENTIAL.
FT	DOMAIN	461	897	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	129	238	FIBROECTIN TYPE-III.
FT	DOMAIN	336	434	FIBROECTIN TYPE-III.
FT	DISULFID	35	45	BY SIMILARITY.
FT	DISULFID	75	91	BY SIMILARITY.
FT	CARBOHYD	58	58	POTENTIAL.
FT	CARBOHYD	191	191	POTENTIAL.
FT	CARBOHYD	346	346	POTENTIAL.
FT	SEQUENCE	897 AA;	97335 MW;	339BE37FDB8F393A CRC64;

Query Match	4.3%	Score 113;	DB 1;	Length
Best Local Similarity	31.9%;	Pred. No. 8.46e-03;		
Matches	22;	Conservative	10;	Mismatches 34;
				Index

```

Db . 372 FETQYRKDTAT-WKDSKTETIQNAHSMALPALEPSTRWARYVRVTSRT
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 148 YEYQYRSPEDTIEW-QSKQENTCNVTIEGLDA-EKCYSFWRVKAMEDVYI

```

DB 431 ARSWDTESV 439  
| : :  
QY 206 VTCWQRGFI 214

RESULT	15	
ID	CPC7_RAT	
AC	P05179;	STANDARD;
		PRT; 490 AA.

DT	15-AUG-1987	(Rel. 03, Created)
DT	01-OCT-1989	(Rel. 12, Last sequence update)
DT	15-DEC-1998	(Rel. 37, Last annotation update)
DE	CYTCHROME P450 2C7	(EC 1.14.14.1) (CYP11C7) (P450F)

GN  
CIPZC/ OR CIPZC-7.  
OC Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Murinae; Rodentia; Sciurognathi; Muridae; Rattus

RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=LIVER;
RX	MEDLINE; 88087187.

Kimura H., Toshioaka H., Sogawa K., Sakai Y., Fujii-Kuriyama M.  
"Complementary DNA cloning of cytochrome P-450s related  
from the complementary DNA library of female rat liver  
primary structures for P-450f, PB-1, and PB-1-related

RT J. Biol. Chem. 263:701-707(1988).  
RN [2]

RP SEQUENCE OF 8-490 FROM N.A.  
RX MEDLINE; 86278140.  
RA Gonzalez F.J., Kimura S., Song B.-J., Pastewka J., Gelb  
RA Hardwick J.P.;

RT sequence of two related P-450 mRNAs transcriptionally  
RT during rat development. An R.dre.1 sequence occupies the  
RT untranslated region of a liver mRNA.";  
J. Biol. Chem. 261:10667-10672(1986).

SEQUENCE OF 87-490 FROM N.A.  
MEDLINE; 87101095.  
Friedberg T., Waxman D.J., Atchison M., Kumar A., Haas

RA Raphael C., Adesnik M.;  
 RT "Isolation and characterization of cDNA clones for cytochrome P-450 2B10  
 RT RT: immunogenically related to rat hepatic P-450 form PB-1  
 RL Biochemistry 25:7975-7983(1986)

1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOL  
MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED  
NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES  
OF STRUCTURALLY UNRELATED COMPOUNDS INCLUDING SMER

CCC	ACIDS, AND XENOBIOTICS.
CCC	-1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O <sub>2</sub>
CCC	-1- OXIDIZED FLAVOPROTEIN + H <sub>2</sub> O.
CCC	-1- SURFICIAL LOCATION: MEMBRANE-BOUND ENDOPLASMIC

CC -!- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER  
CC TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS



US-09-376-430-2-02.rsp

Thu May 11 06:49:28 2000

```

CC      AND CARCINOGENS.
CC      -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M31031; AAA41058.1; -.
CC      EMBL; M18335; AAA41036.1; -.
CC      PIR; A25585; A25585.
CC      PIR; B28516; B28516.
CC      PIR; B25954; B25954.
CC      PRINTS; PR00385; P450.
CC      PRINTS; PR00463; EP450I.
CC      PROSITE; PS00086; CYTOCHROME_P450; 1.
CC      PFAM; PF00067; P450; 1.
CC      Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
CC      Microsome; Endoplasmic reticulum.
CC      BINDING; 435 435
CC      HEME.
CC      K -> E (IN REF. 2).
CC      L -> V (IN REF. 3).
CC      N -> I (IN REF. 2 AND 3).
CC      O -> T (IN REF. 3).
CC      Q -> H (IN REF. 2).
CC      MI -> HD (IN REF. 2).
CC      KGT -> RRA (IN REF. 3).
CC      DPGH -> VPWF (IN REF. 3).
CC      P -> A (IN REF. 3).
CC      P -> A (IN REF. 3).
CC      SEQUENCE 490 AA; 56187 MW; 0FEACF3E9C6ABAA5 CRC64;

```

```

Query Match      4.2%; Score 108; DB 1; Length 490;
Best Local Similarity 32.7%; Pred.No. 4.45e-02;
Matches 18; Conservative 17; Mismatches 18; Indels 2; Gaps 2;

```

```

Db      3 LVTFLL-VLTSLLSLLSLWRQSRRLK-L-PGGTPTLPPIGNFLQIDVKNISQSLT 55
QY      LSSLAAILMVSLLLSLWKLWRKFLIPSVDPKSIFFPGLFEIHQGNFQEWIT 288

```

Search completed: Wed May 10 11:13:14 2000  
Job time : 87 secs.



\*\*\*\*\*  
W P S R L H  
\*\*\*\*\*  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:13:33 2000; Maspar time 194.33 Seconds  
Tabular output not generated. 124.518 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (23-371) from US09376430A.ppt (2 of 25)  
Perfect score: 2601  
Sequence: 1 QGGAAGVQIIIFNLETV.....DVTIGGFVMDRSYVAL 349

Scoring table: PAM 150  
Gap 11

Searched: 225878 seqs, 69334122 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: spiremb12  
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
9:sp-phase 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus

Statistics: Mean 47.400; Variance 87.339; scale 0.543

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	178	6.8	467	11	GFI-2.	6.31e-13
2	130	5.0	383	11	IL-13 RECEPTOR ALPHA 2	6.25e-05
3	114	4.4	422	2	HYPOTHETICAL 45.4 KD P	1.54e-02
4	110	4.2	66	11	CYTOKROME P-450.	5.72e-02
5	110	4.2	326	2	YOP TRANSLOCATION J.	1.09e-01
6	108	4.2	427	4	INTERLEUKIN-13 RECEPTOR	7.90e-02
7	109	4.2	435	2	PUTATIVE ASPARTATE AMI	2.06e-01
8	106	4.1	435	2	ASPARTATE AMINOTRANSFERASE	2.06e-01
9	106	4.1	108	11	ALPHA 7C INTEGRIN (FRA	2.06e-01
10	106	4.1	722	10	PREDICTED PROTEIN OF U	1.50e-01
11	107	4.1	787	11	INTEGRIN BETA-6 SUBUNI	2.06e-01
12	106	4.1	1106	11	INTEGRIN SUBUNIT ALPHA	2.06e-01
13	106	4.1	1167	4	INTEGRIN ALPHA 7 PRECU	1.50e-01
14	106	4.1	1180	11	INTEGRIN ALPHA 7 PRECU	1.50e-01
15	107	4.1	1422	5	PF5NF2L.	2.82e-01
16	105	4.0	359	11	C-C CHEMOKINE RECEPTOR	2.82e-01
17	104	4.0	634	5	SOMATOSTATIN RECEPTOR	3.85e-01
18	104	4.0	1049	5	PF5N8.1 PROTEIN.	3.85e-01
19	104	4.0	1135	11	INTERLEUKIN-13 RECEPTOR	9.72e-01
20	101	3.9	341	13	INTERLEUKIN-10 RECEPTOR	9.72e-01

21	101	3.9	359	6	018863	INTERPHOTORECEPTOR RET	9.72e-01
22	102	3.9	365	5	061833	C23H5.1 PROTEIN.	7.16e-01
23	101	3.9	371	14	091F8	LATENT MEMBRANE PROTEI	9.72e-01
24	101	3.9	381	14	089558	LATENT MEMBRANE PROTEI	9.72e-01
25	102	3.9	415	11	0920K4	INTERLEUKIN-5 RECEPTOR	9.72e-01
26	101	3.9	418	2	032261	YVFL PROTEIN.	9.72e-01
27	101	3.9	421	2	007010	HYPOTHETICAL 47.3 KD P	9.72e-01
28	102	3.9	638	6	09XS21	GROWTH HORMONE RECEPTO	7.16e-01
29	102	3.9	767	5	020170	F3B11.7 PROTEIN.	7.16e-01
30	102	3.9	861	2	006944	HYPOTHETICAL 96.7 KD P	1.32e+00
31	100	3.8	169	2	035628	HYPOTHETICAL 18.4 KD P	1.32e+00
32	100	3.8	183	11	084107	RECEPTOR TYROSINE KINA	1.32e+00
33	100	3.8	220	11	084108	RECEPTOR TYROSINE KINA	1.32e+00
34	100	3.8	233	11	0822M0	INTEGRIN ALPHA IIB SUB	1.32e+00
35	99	3.8	371	14	086784	LATENT MEMBRANE PROTEI	1.78e+00
36	99	3.8	382	14	051FNS	LATENT MEMBRANE PROTEI	1.78e+00
37	99	3.8	382	14	086852	LATENT MEMBRANE PROTEI	1.78e+00
38	99	3.8	386	14	083388	LATENT MEMBRANE PROTEI	1.78e+00
39	100	3.8	506	3	043053	GLUCOSYLTRANSFERASE.	1.32e+00
40	100	3.8	864	13	073637	PHEROMONE RECEPTOR.	1.32e+00
41	100	3.8	890	11	0921A0	INTERLEUKIN-5 RECEPTOR	1.32e+00
42	100	3.8	911	11	035407	RECEPTOR-LIKE TYROSINE	1.32e+00
43	99	3.8	4436	1	058659	4436AA LONG HYPOTHETIC	1.78e+00
44	99	3.8	6048	5	023020	TWITCHIN.	1.78e+00
45	99	3.8	7160	5	023551	ZK617.1B PROTEIN.	1.78e+00

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	467 AA.
ID	Q63216			
AC	Q63216			
DT	01-NOV-1996	(TREMBlrel. 01, Created)		
DT	01-NOV-1996	(TREMBlrel. 01, Last sequence update)		
DT	01-NOV-1998	(TREMBlrel. 08, Last annotation update)		
DE	GFI-2.			
GN	GFI-2.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 95055995.			
RA	FLUBACHER M.M., BEAR S.E., TSICHLIS P.N.;			
RT	"Replacement of interleukin-2 (IL-2)-generated mitogenic signals by a			
RT	mink cell focus-forming (MCF) or xenotropic virus-induced IL-9-			
RT	dependent autocrine loop: implications for MCF virus-induced			
RT	leukemogenesis."			
RL	J. Virol. 68:7709-7716(1994).			
DR	EMBL; L36459; AAA63702.1;			
SQ	SEQUENCE 467 AA; 52017 MW; 0F3E477F CRC32;			
Query Match 6.8%; Score 178; DB 11; Length 467;				
Best Local Similarity 30.08; Pred. No. 6.31e-13;				
Matches 24; Conservative 24; Mismatches 29; Indels 3; Gaps 3;				
Db	272	ILVAVFILLTGLHIF-LFRLSPKVKRIFYQNVPSAEFFHLYSVHGDFTWIGARR 330		
QY	233	ILSSIAILLMVSLLLSLWKLW-RVKKFLIPSPDPKSIFFGLFEIHQGNFOWEITDQ 291		
Db	331	AGQAOQ-DCASAPSGDSSES 349		
QY	292	NVAHLHKMAGAEQESGPEEP 311		
PRELIMINARY: PRT: 383 AA.				
RESULT	2			
ID	Q88786			
AC	Q88786			
DT	01-NOV-1998	(TREMBlrel. 08, Created)		
DT	01-NOV-1998	(TREMBlrel. 08, Last sequence update)		
DT	01-NOV-1998	(TREMBlrel. 08, Last annotation update)		
DE	IL-13 RECEPTOR ALPHA 2.			
OS	Mus musculus (Mouse).			

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]

RC STRAIN-C3H/HEJ; TISSUE=THYMUS;  
 RX MEDLINE; 98391042.  
 RA DONALDSON D.D., WHITTERS M.J., FITZ L., NEBEN T.Y., FINNERTY H.,  
 RA HENDERSON S.L., O'HARA R.M. JR., BEIER D.R., TURNER K.J., WOOD C.R.,  
 RA COLLINS M.;  
 RT "The murine IL-13 receptor alpha 2: molecular cloning,  
 RT characterization, and comparison with murine IL-13 receptor alpha 1.";  
 RL J. Immunol. 161:2317-2324(1998).  
 DR EMBL; U65747; AAC33240.1;  
 SQ SEQUENCE 383 AA; 44483 MW; 5EAEF3E3 CRC32;

Query Match 5.0%; Score 130; DB 11; Length 383;  
 Best Local Similarity 30.2%; Pred. No. 6.25e-05;  
 Matches 26; Conservative 22; Mismatches 32; Indels 6; Gaps 5;

Db 127 EGSLETKIQDMKCIYYNWOVLVCSWPKGVYSDNTYTMFFWEGLDHAL-QCADDYLQHD 185  
 QY 23 QGGAAGVQ-IQIYFLETFVQVTWNSK--YSTRNLTFHYRFNG-DEAYDQCTNVLQEQ 78  
 Db 186 EKNVGCKLSLSDSYKDFICVNGS 211  
 QY 79 GHTSGCLLDAEQRDDIL-YFSIRNGT 103

RESULT 3  
 ID O86640 PRELIMINARY; PRT; 422 AA.  
 AC O86640;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DE HYPOTHETICAL 45.4 KD PROTEIN.  
 GN SC3C3.06C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.  
 RN [1]  
 RC STRAIN-A3(2);  
 RA MURPHY L., HARRIS D.;  
 RA Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RC STRAIN-A3(2);  
 RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RC STRAIN-A3(2);  
 RX MEDLINE; 97000351.  
 RA KINASHI H., HOPWOOD D.A.;  
 RA REDENBACH M., KIESER H.M., DENAPATE D., EICHNER A., CULLUM J.,  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL; AL031231; CAA20255.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 422 AA; 45397 MW; AS27C143 CRC32;

Query Match 4.4%; Score 114; DB 2; Length 422;  
 Best Local Similarity 26.2%; Pred. No. 1.54e-02;  
 Matches 22; Conservative 23; Mismatches 36; Indels 3; Gaps 3;

Db 32 SRLQRLVLAALALGVGVVAVPAALLAVALVIAFARRGRSVPDWLSTARGL 91  
 QY 236 SSLALLMVSLLLSLWKLVRKFF-LIPSPDPKRSIFPGLFEIHQG-NFOEWITDQNV 293  
 Db 92 RHRURRAGVAIPPGTEPALVPAL 115  
 QY 294 AH-LHKMAGAEQESGPEEPLVQL 316

RESULT 4  
 ID Q63706 PRELIMINARY; PRT; 66 AA.  
 AC Q63706;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DE CYTOCHROME P-450.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RC STRAIN-SPRAGUE-DAWLEY;  
 RA STROEM A., NILSSON A.G., ZAPHIROPOULOS P.;  
 RL Nucleic Acids Res. 0:0-0(1988).  
 DR EMBL; X12595; CAA31108.1;  
 SQ SEQUENCE 66 AA; 7298 MW; CF36233A CRC32;

Query Match 4.2%; Score 110; DB 11; Length 66;  
 Best Local Similarity 31.1%; Pred. No. 5.72e-02;  
 Matches 19; Conservative 18; Mismatches 22; Indels 2; Gaps 2;

Db 3 LVTFLL-VLATSSLLLSLWRSRRKL-PPGPTPLPIIGNFLQIDVKNISQSLTKVSIV 60  
 QY 234 LISSLAILLMVSLLLSLWKLVRKFLIPSPDKSIFPGLFEIHQGNFQEWITDQNV 293

Db 61 G 61  
 QY 294 A 294

RESULT 5  
 ID O84563 PRELIMINARY; PRT; 326 AA.  
 AC O84563;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE YOP TRANSLLOCATION J.  
 GN YSCJ.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 RN [1]  
 RC STRAIN-D/UW-3/CX;  
 RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L.,  
 RA MITCHELL W.P., OLINGER L., TATSOV R.L., ZHAO Q., KOONIN E.V.,  
 RA DAVIS R.W.;  
 RT "Genome Sequence of an Obligate Intracellular Pathogen of Humans:  
 RT Chlamydia trachomatis";  
 RL Science 0:0-0(1998).  
 RN [2]  
 RC STRAIN-D/UW-3/CX;  
 RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L.,  
 RA MITCHELL W.P., OLINGER L., TATSOV R.L., ZHAO Q., KOONIN E.V.,  
 RA DAVIS R.W.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE001327; AAC68161.1;  
 DR PFAM; PF01514; YSCJ\_FliF; 1.  
 SQ SEQUENCE 326 AA; 35553 MW; 0B76FEA3 CRC32;

Query Match 4.2%; Score 110; DB 2; Length 326;  
 Best Local Similarity 32.1%; Pred. No. 5.72e-02;  
 Matches 17; Conservative 14; Mismatches 20; Indels 2; Gaps 2;

Db 229 AKHSLTRFLRVFVFLILLFLLSCGLLWVTKHT-LISALGKTGFDPAPY 280  
 QY 225 PKKLSKFLISSLAILLMVSLLLSLWKLVRKFLIPSPDKSIF-PGLF 276

RESULT 6

```
ID O95646 PRELIMINARY; PRT; 427 AA.
AC O95646;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DE 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE INTERLEUKIN-13 RECEPTOR.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RN SEQUENCE FROM N.A.
RA WADA M., HISANO T., KOWANO M.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81379; AAD00510.2; -.
KW RECEPTOR.
SQ SEQUENCE 427 AA; 48737 MW; 52BF0F51 CRC32;

Query Match 4.2%; Score 108; DB 4; Length 427;
Best Local Similarity 22.5%; Pred. No. 1.09e-01;
Matches 64; Conservative 71; Mismatches 123; Indels 26; Gaps 26;

Db 122 EGDPSAVTELCQIWHNLSYMKCSWLPGRNTPDNTLYYWHRSLEKIHCEN-IFREG 180
QY 23 QGGRAEGV-QIQLIYFNLETVQVITWNSKY-SR-TNLTFFHYRFNGDEAYDQCTNYLLQEG 79
Db 181 QYFGCSPLTKVQSSPQHSVQIMVKDNAGKIRPSFNIIP-LTSRVKPDPP-HIKNLSF 238
QY 80 HTSGC-L-LD-A-EQR-DD-ILYFSIR-N-GT-HEVFTASRWVYLLKPSPKHVR-FSW 129
Db 239 HNDLLYQWENPQNFYSCLFYEYEVNNSQTHNVFVQAKENPEFERNVENTSCFM 298
QY 130 HODAVTVTCSD-LS-IGD-LIFEVO-YRSPDTEWQS-KQENTCNVTIEGLDAEKCYSEW 184
Db 299 VPG-VLPDLTNVTRVK-TNKLVEDKLNSWQSEMSIGKKNSTLYITMLLPVIV 356
QY 185 VYVKAMEDVYGDTPVPSWSEVTCWQGEIRDA-CAETPTPKPKLSKFI-LISSIALLL 242.
Db 357 AGAIVLLY-LKRLKIIFPPDPGKIFKEMFGDQNDTLHW 399
QY 243 MVSLLLSLWLKRVYKFLIPSPDPKSIFFGLFEIHOGNFOEW 286

RESULT 7
ID O06685 PRELIMINARY; PRT; 435 AA.
AC O06685;
DT 01-JUL-1997 (TREMblrel. 04, Created)
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DE 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE PUTATIVE ASPARTATE AMINOTRANSFERASE TPAAT.
CN TPAAT.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
RN [1]
RN SEQUENCE FROM N.A.
RA SHEVCHENKO D.V., AKINS D.R., ROBINSON E., LI M., POPOVA T.G.;
RX MEDLINE; 97294457.
RA COX D.L., RADOLF J.;
RT "Molecular characterization and cellular localization of TplRR, a
RT processed leucine-rich repeat protein of Treponema pallidum, the
RT syphilis spirochete."
RL J. Bacteriol. 179:3188-3195(1997).
DR EMBL; U73748; AAC45303.1; -.
KW Transferase; Aminotransferase.
SQ SEQUENCE 435 AA; 47285 MW; 7AF01306 CRC32;

Query Match 4.2%; Score 109; DB 2; Length 435;
Best Local Similarity 35.1%; Pred. No. 7.90e-02;
Matches 20; Conservative 11; Mismatches 21; Indels 5; Gaps 5;

Db 206 AVLVICDD-AYSGFEYASLMRGSEFAFAQAQKNCALKIDGLTKEE-YA-WGLRV 259
QY 133 AVTVTCSDLSYGDLLYEVOY-RSPFDTEWQSKQENTCNVTIEGLDAEKCYSEW-VRV 187

RESULT 9
ID Q63027 PRELIMINARY; PRT; 108 AA.
AC Q63027;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE ALPHA 7C INTEGRIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RN SEQUENCE FROM N.A.
RA SONG W.K., WANG W., SATO H., BIELSER D., KAUFMAN S.;
RX MEDLINE; 94171924.
RA "Expression of alpha 7 integrin cytoplasmic domains during skeletal
RT muscle development: alternate forms, conformational change, and
RT homologies with serine/threonine kinases and tyrosine phosphatases."
RL Cell Sci. 106:1139-1152(1993).
DR EMBL; X74294; CAA52347.1; -.
DR PFAM; PF00357; integrin_A; 1.
KW Integrin.
FT NON-TER
SQ SEQUENCE 108 AA; 12016 MW; 1A077241 CRC32;
```

```
RESULT 8
ID O83252 PRELIMINARY; PRT; 435 AA.
AC O83252;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DE 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE ASPARTATE AMINOTRANSFERASE (TPAAT).
CN TPAAT.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
RN [1]
RN SEQUENCE FROM N.A.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.;
RX MEDLINE; 98332770.
RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.;
RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.;
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.;
RA MCDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.;
RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.;
RA VENTER J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
RN [2]
RN SEQUENCE FROM N.A.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.;
RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.;
RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.;
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.;
RA MCDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.;
RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.;
RA VENTER J.C.;
RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE001204; AAC65212.1; -.
RX TIGR; TP0223.
KW Transferase; Aminotransferase.
SQ SEQUENCE 435 AA; 47304 MW; 0B31A6E1 CRC32;

Query Match 4.2%; Score 109; DB 2; Length 435;
Best Local Similarity 35.1%; Pred. No. 7.90e-02;
Matches 20; Conservative 11; Mismatches 21; Indels 5; Gaps 5;

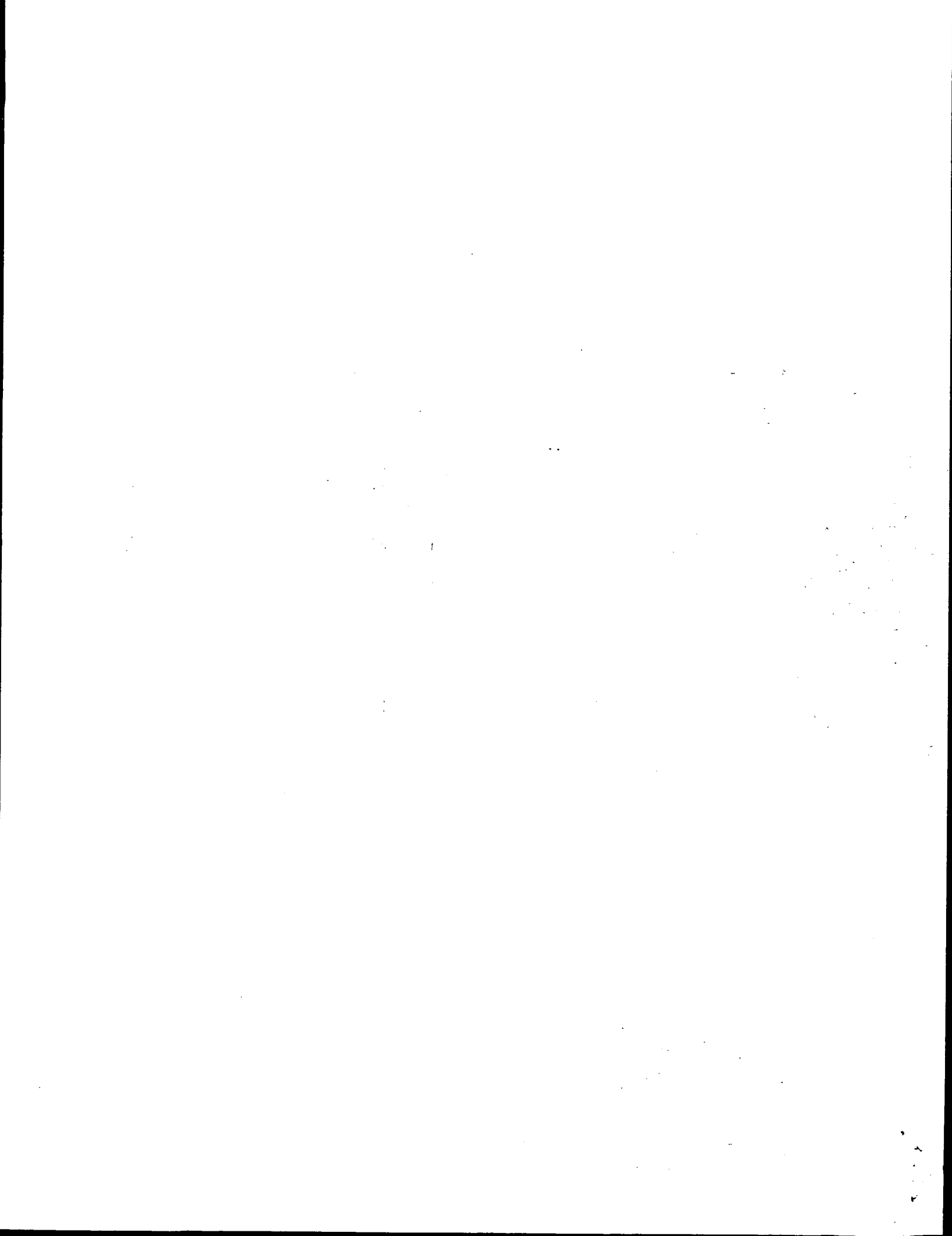
Db 206 AVLVICDD-AYSGFEYASLMRGSEFAFAQAQKNCALKIDGLTKEE-YA-WGLRV 259
QY 133 AVTVTCSDLSYGDLLYEVOY-RSPFDTEWQSKQENTCNVTIEGLDAEKCYSEW-VRV 187

RESULT 9
ID Q63027 PRELIMINARY; PRT; 108 AA.
AC Q63027;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE ALPHA 7C INTEGRIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RN SEQUENCE FROM N.A.
RA SONG W.K., WANG W., SATO H., BIELSER D., KAUFMAN S.;
RX MEDLINE; 94171924.
RA "Expression of alpha 7 integrin cytoplasmic domains during skeletal
RT muscle development: alternate forms, conformational change, and
RT homologies with serine/threonine kinases and tyrosine phosphatases."
RL Cell Sci. 106:1139-1152(1993).
DR EMBL; X74294; CAA52347.1; -.
DR PFAM; PF00357; integrin_A; 1.
KW Integrin.
FT NON-TER
SQ SEQUENCE 108 AA; 12016 MW; 1A077241 CRC32;
```



SQ SEQUENCE 1167 AA; 127574 MW; 8290E63E CRC32;  
Query Match 4.1%; Score 106; DB 4; Length 1167;  
Best Local Similarity 51.6%; Pred. No. 2.06e-01;  
Matches 16; Conservative 8; Mismatches 5; Indels 2; Gaps 2;  
DB 1117 RPLISLWTLIGSVGGILLALLVFCWLKL 1147  
QY 226 KPLK-SKFLISS-LALLMVSLLLSLWL 254  
RESULT 14  
ID O88732 PRELIMINARY; PRT: 1180 AA.  
AC O88732;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
DE INTEGRIN ALPHA 7 PRECURSOR.  
GN ITGA7.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129SV;  
RA SAHPR G., ECHTERMEYER F., BEIER D.R., POESCHL E., MAYER U.;  
RT "Genomic organization and chromosomal localization of the mouse  
integrated alpha7 gene";  
RL Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Y12380; CAA73023.1; JOINED.  
DR EMBL; Y12383; CAA73023.1; JOINED.  
DR EMBL; Y12384; CAA73023.1; JOINED.  
DR EMBL; Y12385; CAA73023.1; JOINED.  
DR EMBL; Y12386; CAA73023.1; JOINED.  
DR EMBL; Y12387; CAA73023.1; JOINED.  
DR EMBL; Y12388; CAA73023.1; JOINED.  
DR EMBL; Y12389; CAA73023.1; JOINED.  
DR EMBL; Y12390; CAA73023.1; JOINED.  
DR EMBL; Y12392; CAA73023.1; JOINED.  
DR HSSP; P11215; IABX.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PFAM; PF00357; Integrin\_A; 3.  
KW Signal; Integrin.  
FT SIGNAL 0 32 POTENTIAL.  
SQ SEQUENCE 1180 AA; 129261 MW; 57E6BE73 CRC32;  
Query Match 4.1%; Score 106; DB 11; Length 1180;  
Best Local Similarity 59.1%; Pred. No. 2.06e-01;  
Matches 13; Conservative 6; Mismatches 3; Indels 0; Gaps 0;  
DB 1084 ILLAVLAGLVALLVLLWLKL 1105  
QY 233 ILISSLAILLMVSLLLSLWL 254  
RESULT 15  
ID O00914 PRELIMINARY; PRT: 1422 AA.  
AC O00914;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE PFSNF2L.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97418900.  
RA JI D.D., ARNOT D.E.;  
RT "A Plasmodium falciparum homologue of the ATPase subunit of a multi-  
protein complex involved in chromatin remodelling for transcription.";  
RL Mol. Biochem. Parasitol. 88:151-162(1997).  
DR EMBL; AF003086; AAC47719.1; -.  
DR PFAM; PF00271; helicase\_C; 1.

DR PFAM; PF00176; SNF2\_N; 1.  
SQ SEQUENCE 1422 AA; 166947 MW; FD02C912 CRC32;  
Query Match 4.1%; Score 107; DB 5; Length 1422;  
Best Local Similarity 26.1%; Pred. No. 1.50e-01;  
Matches 18; Conservative 19; Mismatches 30; Indels 2; Gaps 2;  
DB 463 RF-LRSENRLITGTPLHNNKELWLLNFKLPKIFDNSEEDNLFNISKISTND-NKQS 520  
QY 231 KFLISSLAILLMVSLLLSLWLKLVKFLIPSVDPKSIFFGLFEIHQGNFQEWITDT 290  
DB 521 ELITOLHTI 529  
QY 291 QNVHLHKM 299  
Search completed: Wed May 10 11:16:57 2000  
Job time : 204 secs.





\*\*\*\*\*  
 WQSEH  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 11:26:42 2000; MasPar time 9.68 Seconds  
 Tabular output not generated. 565.188 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (1-231) from US09376430A.pap (3 of 25)  
 Perfect Score: 1773  
 Sequence: 1 MGRLLVLLGAAVFLGGWMA.....GEIRDACAEETPPKPKLSK 231

Scoring table: PAM 150  
 Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: a-genseq35  
 1:geneseq

Statistics: Mean 32.727; Variance 127.288; scale 0.257

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	200	11.3	482	1 W31646	Human cytokine recepto	4.24e-09
2	195	11.0	230	1 R82934	Interleukin 4 componen	1.28e-08
3	195	11.0	230	1 R47151	IL-2 receptor gamma ch	1.28e-08
4	195	11.0	252	1 R47150	IL-2 receptor gamma ch	1.28e-08
5	195	11.0	347	1 R47149	IL-2 receptor gamma ch	1.28e-08
6	195	11.0	369	1 R47148	IL-2 receptor gamma ch	1.28e-08
7	188	10.6	369	1 R5094	Murine IL-2R gamma.	5.92e-08
8	132	7.4	383	1 W35294	Murine IL-13 binding c	8.30e-03
9	115	6.5	576	1 R78613	Expression vector pME1	2.45e-01
10	115	6.5	596	1 R78616	Expression vector pME1	2.45e-01
11	115	6.5	600	1 R92536	Fas antigen #1.	2.45e-01
12	115	6.5	600	1 R78610	Expression vector pME1	2.45e-01
13	113	6.4	897	1 R78608	Murine IL-3 receptor b	2.45e-01
14	111	6.3	227	1 R20982	Sequence of beta-chain	3.62e-01
15	111	6.3	237	1 R22229	Truncated human growth	5.34e-01
16	111	6.3	269	1 R05045	Soluble human growth h	5.34e-01
17	112	6.3	292	1 R92527	Fas antigen #2.	4.40e-01
18	110	6.2	237	1 R24274	Truncated human growth	6.47e-01
19	109	6.1	878	1 R92529	Fas sequence from AIC2	7.85e-01
20	105	5.9	572	1 Y04954	Mycobacterium species	1.68e-00
21	102	5.8	315	1 W56261	Mature interleukin-13	2.97e+00
22	102	5.8	359	1 W56260	Construct containing m	2.97e+00
23	102	5.8	380	1 W35295	Human IL-13 binding ch	2.97e+00

24	102	5.8	380	1 W33603	Homo sapiens HR-1 rece	2.97e+00
25	102	5.8	380	1 W36613	Human zcytor2 cytokine	2.97e+00
26	102	5.8	380	1 W24972	Human interleukin-13 b	2.97e+00
27	102	5.8	380	1 W41502	Human cytokine/peptide	2.97e+00
28	102	5.8	380	1 W41520	Human HR-1 receptor.	2.97e+00
29	100	5.6	110	1 R56380	Extracellular hGR-CD,	4.31e+00
30	100	5.6	112	1 R56391	Human growth hormone r	4.31e+00
31	100	5.6	246	1 R56389	Human growth hormone r	4.31e+00
32	100	5.6	249	1 W0426	Hormone binding region	4.31e+00
33	100	5.6	269	1 R06867	Human somatogenic rece	4.31e+00
34	100	5.6	380	1 W36614	Human zcytor2 cytokine	4.31e+00
35	100	5.6	637	1 P92108	Human growth hormone r	4.31e+00
36	100	5.6	638	1 W33394	Human growth hormone r	4.31e+00
37	100	5.6	638	1 P81326	Human growth hormone r	4.31e+00
38	97	5.5	372	1 W36616	Celebus macaque zcytor	7.52e+00
39	94	5.3	269	1 W82802	Human soluble somatoge	1.30e+01
40	94	5.3	638	1 P92107	Rabbit growth hormone	1.30e+01
41	94	5.3	638	1 P81327	Rabbit growth hormone	1.30e+01
42	94	5.3	638	1 W33395	Rabbit growth hormone	1.30e+01
43	90	5.1	426	1 W09822	Human interleukin-12 r	2.89e+01
44	90	5.1	427	1 W24973	Human interleukin-13 a	2.89e+01
45	90	5.1	1239	1 R45945	Glutamic acid receptor	2.69e+01

ALIGNMENTS

RESULT 1  
 ID W31646 standard; Protein; 482 AA.  
 AC W31646;  
 DT 21-MAY-1998 (first entry)  
 DE Human cytokine receptor gc chain-Ig fusion protein.  
 KW Cytokine receptor; gamma common chain; gc chain; human;  
 KW blocking agent; monoclonal antibody; CP.B8; immunological disease;  
 KW myasthenia gravis; rheumatoid arthritis; lupus; multiple sclerosis;  
 KW insulin-dependent diabetes; inflammatory bowel disease;  
 KW sympathetic ophthalmia; uveitis; allergy; asthma; infection;  
 KW graft versus host disease; psoriasis; immunosuppressive; therapy.  
 OS Chimeric - Homo sapiens.  
 FH Key  
 FT Location/Qualifiers  
 FT 1..254  
 FT /note="gc chain N-terminal region"  
 FT Protein  
 FT 255..482  
 FT /note="IgG1 constant region"  
 FT Region  
 FT 255..264  
 FT /note="IgG1 hinge region"  
 FT Domain  
 FT 284..482  
 FT /note="IgG1 CH2 and CH3 constant domains1"  
 PN WC09743416-A1.  
 PD 20-NOV-1997.  
 PF 09-MAY-1997; U07870.  
 PR 10-MAY-1996; US-017466.  
 PA (BIOU ) BIOGEN INC.  
 PI Benjamin CD, Burkiy LC, Hession C, Whitty A;  
 DR WPI: 98-00885/01.  
 DR N-PSDB: T97439.  
 PT Blocking agents of the gamma common chain of cytokine receptors -  
 PT treatment of immunological diseases  
 PS Example 1; Page 79-80; 111pp; English.  
 CC This polypeptide comprises a fusion between the N-terminal 254  
 CC amino acids of the human mature cytokine receptor gamma common (gc)  
 CC chain and the hinge region and CH2 and CH3 constant domains of  
 CC human IgG1. The fusion was expressed from clone pLB001 (see  
 CC T97439) in COS-7 cells, and used to generate murine anti-human gc  
 CC specific monoclonal antibodies (MAbs), including CP.B8 produced by  
 CC hybridoma ATCC HB 12107. The invention provides compositions and  
 CC methods for inhibiting cytokine signalling using gc chain blocking  
 CC agents for the treatment of immunological diseases such as  
 CC myasthenia gravis, rheumatoid arthritis, lupus, multiple sclerosis,  
 CC insulin-dependent diabetes, inflammatory bowel disease, sympathetic  
 CC ophthalmia, uveitis, allergy, asthma, parasitic infection, graft  
 CC vs. host disease or psoriasis. A preferred gc blocking agent is  
 CC Mab CP.B8 or its Fab fragment (see also W31647-48).

```

SQ Sequence 482 AA;
Query Match 11.3%; Score 200; DB 1; Length 482;
Best Local Similarity 25.7%; Pred. No. 4.24e-09;
Matches 53; Conservative 53; Mismatches 88; Indels 12; Gaps 12;

Db 59 EVQCFVFNVEYMNCTWNSSESPQPTNLTHYWKNSDNDKVKCHSHYLFSEITSGCQLQ 118
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :
Qy 31 QIQIIFNLETVOVTWNASKYSR-TNLTFRYF-NGD-EAYDQCTNLLQEGHTSGCLLD 87
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :

Db 119 KKEIHLXYOTFVVQLQDPREPRQATQMLKQNLVTPWAPENLTLHKLSOLELNNWRF 178
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :
Qy 88 AEQRDILYFSIR-NGTH-PVFTASRW-VYYL-KPSSPKHVRP-SWHQDAVTVTCSCLS 142
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :

Db 179 LNHCLHLYVQYRTDWDHSWTEQSDVYRKFSLPVSDGQKRYFRVRSR-FNPLCGSAQHW 237
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :
Qy 143 YGD-LLEYEQYRSPDTEW-QSKQENTCNVTIEGLDAEKCYSFWRVKAMEDVYGPDTYP 200
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :

Db 238 SEWSPHPIHWSNTSKENVDKHTHTCPP 263
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :
Qy 201 SDWSEVTCWQGEIRDACAETPT-PP 225
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :

RESULT 2
ID R82934 standard; Protein; 230 AA.
AC R82934;
DE Interleukin-4 component common to the IL-2 receptor gamma chain.
KW Interleukin-4; IL-4; gamma chain component; immunosuppressants;
KW anti-allergy agent; signal transmission inhibitor; autoimmune;
KW disease; anti-inflammatory; anaphylactic shock; bronchial asthma;
KW interleukin-2; IL-2; atopic dermatitis; urticaria.
OS Homo sapiens.
PN J07149662-A.
PD 13-JUN-1995.
PF 07-SEP-1994; 213706.
PR 08-SEP-1993; JP-223374.
PA (AJIN ) AJINOMOTO KK.
PA (SUGA/) SUGAMURA K.
DR WPI; 95-243601/32.
DR N-PSDB; T04952.
DR Novel interleukin-4 receptor monoclonal antibodies inhibit signal
PT transmission - useful as immunosuppressants and anti-allergy agents.
PS Example 1; Page 9; ilpp: Japanese.
CC T04952 encodes R82934 a component of the IL-4 receptor common to
CC the IL-2 receptor gamma chain molecule, which was used to generate
CC anti-IL-4 receptor monoclonal antibodies (mAbs). The mAbs (IL-4
CC signal transmission inhibitors) can be used as immunosuppressants
CC and anti-allergy agents, for the treatment of autoimmune and chronic
CC inflammatory diseases, e.g. anaphylactic shock, bronchial asthma,
CC atopic dermatitis and urticaria.
SQ Sequence 230 AA;

Query Match 11.0%; Score 195; DB 1; Length 230;
Best Local Similarity 25.9%; Pred. No. 1.28e-08;
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;

Db 38 EVQCFVFNVEYMNCTWNSSESPQPTNLTHYWKNSDNDKVKCHSHYLFSEITSGCQLQ 97
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :
Qy 31 QIQIIFNLETVOVTWNASKYSR-TNLTFRYF-NGD-EAYDQCTNLLQEGHTSGCLLD 87
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :

Db 98 KKEIHLXYOTFVVQLQDPREPRQATQMLKQNLVTPWAPENLTLHKLSOLELNNWRF 157
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :
Qy 88 AEQRDILYFSIR-NGTH-PVFTASRW-VYYL-KPSSPKHVRP-SWHQDAVTVTCSCLS 142
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :

Db 158 LNHCLHLYVQYRTDWDHSWTEQSDVYRKFSLPVSDGQKRYFRVRSR-FNPLCGSAQHW 216
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :
Qy 143 YGD-LLEYEQYRSPDTEW-QSKQENTCNVTIEGLDAEKCYSFWRVKAMEDVYGPDTYP 200
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :

Db 217 SEWSPHPIHW 225
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :
Qy 201 SDWSEVTCW 209
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :

```

```

RESULT 3
ID R47151 standard; Protein; 230 AA.
AC R47151;
DT 13-JUN-1994 (first entry)
DE IL-2 receptor gamma chain.
KW Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;
KW rheumatoid arthritis; transplant rejection; primer;
KW polymerase chain reaction; PCR; amplification.
OS Homo sapiens.
PN Ep-578932-A.
PD 19-JAN-1994.
PF 22-APR-1993; 106561.
PR 23-APR-1992; JP-104947.
PA (AJIN ) AJINOMOTO KK.
PA (SUGA/) SUGAMURA K.
PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;
PI Suzuki M, Takeshita T;
DR WPI; 94-017546/03.
DR P-PSDB; Q54831.
DR DNA and protein sequences of IL-2 gamma chain - useful as immune
PT regulatory agents for treatment of e.g. rheumatoid arthritis and
PT transplant rejection
PS Disclosure; Page 22-23, 35-36; 50pp; English.
CC The human IL-2 receptor gamma chain preform (R47148), including the
CC signal peptide, is encoded by the sequence given in Q54828. The
CC mature protein (R47149) is encoded by sequence Q54829. A soluble
CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,
CC while a soluble form suitable for expression in prokaryotes (R47151)
CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-
CC terminal sequence of IL-2 receptor gamma chain, and are used to
CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27
CC are used to obtain the protein given in R47151.
SQ Sequence 230 AA;

Query Match 11.0%; Score 195; DB 1; Length 230;
Best Local Similarity 25.9%; Pred. No. 1.28e-08;
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;

Db 37 EVQCFVFNVEYMNCTWNSSESPQPTNLTHYWKNSDNDKVKCHSHYLFSEITSGCQLQ 96
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :
Qy 31 QIQIIFNLETVOVTWNASKYSR-TNLTFRYF-NGD-EAYDQCTNLLQEGHTSGCLLD 87
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :

Db 97 KKEIHLXYOTFVVQLQDPREPRQATQMLKQNLVTPWAPENLTLHKLSOLELNNWRF 156
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :
Qy 88 AEQRDILYFSIR-NGTH-PVFTASRW-VYYL-KPSSPKHVRP-SWHQDAVTVTCSCLS 142
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :

Db 157 LNHCLHLYVQYRTDWDHSWTEQSDVYRKFSLPVSDGQKRYFRVRSR-FNPLCGSAQHW 215
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :
Qy 143 YGD-LLEYEQYRSPDTEW-QSKQENTCNVTIEGLDAEKCYSFWRVKAMEDVYGPDTYP 200
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :

Db 216 SEWSPHPIHW 224
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :
Qy 201 SDWSEVTCW 209
   : : | | : | | | : | | | : | | : | | : | | | : | | | : | | | :

RESULT 4
ID R47150 standard; Protein; 252 AA.
AC R47150;
DT 13-JUN-1994 (first entry)
DE IL-2 receptor gamma chain.
KW Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;
KW rheumatoid arthritis; transplant rejection; primer;
KW polymerase chain reaction; PCR; amplification; ss.
OS Homo sapiens.
PN Key Location/Qualifiers
FH Key 1...22 /label= sig_peptide
FT peptide
PN Ep-578932-A.
PD 19-JAN-1994.
PF 22-APR-1993; 106561.
PR 23-APR-1992; JP-104947.
PA (AJIN ) AJINOMOTO KK.

```

Thu May 11 06:49:29 2000

(SUGA/) SUGAMURA K.  
Asao H., Hamuro J., Nakamura M., Shimamura T., Sugamura K;  
Suzuki M., Takeshita T;  
WPI; 94-017546/03.  
P-PSDB: Q54830.  
DNA and protein sequences of IL-2 gamma chain - useful as immune  
regulatory agents for treatment of e.g. rheumatoid arthritis and  
transplant rejection  
Disclosure: Page 21-22, 34-35; 50pp; English.  
The human IL-2 receptor gamma chain preform (R471148), including the  
signal peptide, is encoded by the sequence given in Q54828. The  
mature protein (R47149) is encoded by sequence Q54829. A soluble  
form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
while a soluble form suitable for expression in prokaryotes (R47151)  
is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
terminal sequence of IL-2 receptor gamma chain, and are used to  
isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
are used to obtain the protein given in R47151.  
sequence 252 AA;

Query Match	11.0%	Score	195;	DB 1:	Length	252;			
Best Local Similarity	25.9%	Pred.	Nm.0.128e-08;						
Matches	49;	Conservative	52;	Mismatches	77;	Indels	11;	Gaps	11;
Db	59	EVQCFVFNVEYNCTWNSSEPOPNLTHYWKNSDNDKVKCSHYLSEBITSGCQLO	118						
Qy	31	QIQIYFNLETQVQVTWNAKSYR-TNLTHYR-NGD-EAYDCINLLGEGHTSGCLLD	87						
Db	119	KREIHLQYTFVQLOQPREPQATOMLKLQNLVIPWAPENTLHLKSSOLELWNNRF	178						
Qy	88	AEQRDDILYSIR-NGTH-PVFTSRMM-VYIL-KPSSPKHVF-SWHQDAVTVTCSDL	142						
Db	179	LNHCLHLYQYRTDWSHTEQSDVYRHKFSLEFSDVQGRKYTRVRSR-FNPJCGSAQHW	237						
Qy	143	YGB-LLYEYQYRSPFTEW-QSKQENTCNVTIEGLDAEKYCSFWVRKAMEDYIGPDTYP	200						
Db	238	SEWSHPHW	246						
Qy	201	SDWSEVTCW	209						

RESULT	5	
ID	R47149 standard; Protein; 347 AA.	
AC	R47149;	
DT	13-JUN-1994 (first entry)	
DE	IL-2 receptor gamma chain.	
DE	IL-2 receptor gamma chain;	
DE	Interleukin-2 receptor gamma chain;	
KW	rheumatoid arthritis; transplamt rejection; primer;	
KW	polymerase chain reaction; PCR; amplification.	
KW	homo.sapiens.	
OS	EP-578932-A.	
PN	19-JAN-1994.	
PF	106561.	
PF	22-APR-1993;	
PF	JP-104947.	
PR	(AJIN ) AJINOMOTO KK.	
PA	(SUGA/) SUGAMURA K.	
PI	Asao H, Hamuro J.	
PI	Nakamura M, Shimamura T, Sugamura K;	
PI	Suzuki M, Takeshita T;	
PI	WPI: 94-017545/03.	
DR	P-5PDB; Q54829.	
DR	DNA and protein sequences of IL-2 gamma chain - useful as immune	
PT	regulatory agents for treatment of e.g. rheumatoid arthritis and	
PT	transplant rejection	
PT	Claim 4; Page 41; 50pp; English.	
CC	The human IL-2 receptor gamma chain preform (R47148), including the	
CC	signal peptide, is encoded by the sequence given in Q54828. The	
CC	mature protein (R47149) is encoded by sequence Q54829. A soluble	
CC	form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,	
CC	while a soluble form suitable for expression in prokaryotes (R47151)	
CC	is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-	
CC	terminal sequence of IL-2 receptor gamma chain, and are used to	
CC	isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27	
CC	are used to obtain the protein given in R47151.	
CC		

SQL	Sequence	347 AA:	
	Query Match	11.0%;	Score 195; DB 1; Length 347;
	Best Local Similarity	25.9%;	Pred. No. 1.28e+08;
	Matches	49; Conservative	52; Mismatches 77; Indels 11; Gaps 11;
Db	37	EVQCFVNFVNMCTWNSSEPOPTNLTHYWTYKNSDNKVKQKCHYLSEETISGCOLQ	96
QY	31	QIQIIFYNLETQVTVWNSKYSR-INLTHYRF-NGD-EAYOCTNLYLQEGHTSGCLLD	87
Db	97	KKBIHLIYQTFVVLQDPREPFRQATQMLKLNVLVWAPENLTLLHKLSEQLSEQLNNRNF	156
QY	88	AEQRDDILFYSIR-NGTH-PVFTASRWM-VYYL-KPSSPKHVRF-SWHQDAVTVTCDLS	142
Db	157	LNHCLHLVQRYRDWDHSMTEQSDVYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW	215
QY	143	YGD-LLYEVOGRSPFDTEW-QSKQENTCNVTIEGLDAEKYSFWRVKAMEDVYIGPDTYP	200
Db	216	SEWSHPITHW 224	
QY	201	SDWSEVTCW 209	
RESULT	6		
ID	R47148	standard; Protein; 369 AA.	
AC	R47148;		
DT	13-JUN-1994	(first entry)	
DE	IL-2 receptor gamma chain.		
KW	Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;		
KW	rheumatoid arthritis; transplant rejection; primer; PCR;		
KW	polymerase chain reaction; amplification; ss.		
OS	Homo sapiens.		
FT	Key	Location/Qualifiers	
FT	peptide	1..22	
FT		/label= sig_peptide	
PN	EP-578932-A.		
PD	19-JAN-1994.		
PF	22-APR-1993; 106561.		
PR	23-APR-1992; JP-104947.		
PA	(AJIN) AJINOMOTO KK.		
PA	(SUGA) SUGAMURA K.		
PI	ASAO H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;		
PI	Suzuki M, Takeshita T;		
PI	WPI: 94-017546/03.		
DR	P-SDS; Q54828.		
PT	DNA and protein sequences of IL-2 gamma chain - useful as immune		
PT	regulatory agents for treatment of e.g. rheumatoid arthritis and		
PT	transplant rejection.		
PS	Disclosure; Page 16-17, 29-30; 50pp; English.		
CC	The human IL-2 receptor gamma chain preform (R47148), including the		
CC	signal peptide, is encoded by the sequence given in Q54828. The		
CC	mature protein (R47149) is encoded by sequence Q54829. A soluble		
CC	form of IL-2 receptor gamma chain (R47150) is encoded by Q54830.		
CC	while a soluble form suitable for expression in prokaryotes (R47151)		
CC	is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-		
CC	terminal sequence of IL-2 receptor gamma chain, and are used to		
CC	isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27		
CC	are used to obtain the protein given in R47151.		
SQL	Sequence	369 AA:	

Query Match	11.0%;	Score 195;	DB 1;	Length 369;
Best Local Similarity	25.9%;	Pred. No. 1.28e-08;		
Matches	49;	Conservative 52;	Mismatches 77;	Indels 11; Gaps 11;
Db	59	EVQCFVFNVEYMNCTWNSSSEPOPTNLTHYWKNSDNDKVQCKSHYLFSEITSGCQIQ	118	
Qy	31	QIQIIFNLETVOGTWNASKYSR-TNLTHYRF-NGD-EAYDOCNVLLQEGHTSGCLLD	87	
Db	119	-KKIETHLYQTFVWGLQDPREFRQATQMLKQNLVPIWAPENITLHKLSQLELNNNR	178	
Qy	88	AEQRDILYFSIR-NGTH-PVGTASWM-VYIL-KPSPKHYRF-SWHQDAVIVTCSDL	142	
Db	179	LNCLEHLVOYRTDWDHSMTEQSDVYRHKFSLPSVDGQRKYTFYRSR-FNPLCGSAQH	237	

QY 143 YGD-LLEVOYRSPDTEW-OSKOENTCNVTIEGLDAEKYSEFWVRKAMEDVYGPDTYP 200  
 Db 238 SEWSHPPIHW 246  
 QY 201 SDWSEVTCW 209

RESULT 7  
 ID R59094 standard; Protein; 369 AA.  
 AC R59094;  
 DE Murine IL-2R gamma (first entry)  
 KW Murine IL-2R gamma.  
 KW XSCID; interleukin.  
 OS Mus musculus.  
 FH Key  
 FT Location/Qualifiers  
 FT 1..21  
 FT domain  
 FT 258..284  
 FT /note= "signal peptide"  
 FT /note= "transmembrane domain"  
 FT misc\_difference 331  
 FT /note= "corresponding codon CAG"  
 FT modified\_site 71..73  
 FT /label= N-glycosylation\_site  
 FT modified\_site 75..77  
 FT /label= N-glycosylation\_site  
 FT modified\_site 84..86  
 FT /label= N-glycosylation\_site  
 FT modified\_site 96..98  
 FT /label= N-glycosylation\_site  
 FT modified\_site 159..161  
 FT /label= N-glycosylation\_site  
 FT modified\_site 255..257  
 FT /label= N-glycosylation\_site  
 FT /label= N-glycosylation\_site  
 PN WO9420641-A.  
 PD 15-SEP-1994.  
 PF 10-MAR-1994; U02891.  
 PR 12-MAR-1993; US-031143.  
 PR 14-SEP-1993; US-121435.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PI Leonard WJ, McBride WO, Noguchi M;  
 DR WPI: 94-303046/37.  
 DR N-PSDB: Q19177.

Diagnosis of X-linked severe combined immunodeficiency (XSCID) - comprises detecting mutated IL-2R gamma gene, also vectors and transgenic animals containing the mutated gene  
 PS Example 1; Fig 7; 98pp; English.  
 CC Q19177 is the DNA sequence of murine IL-2R gamma R59094,  
 CC this was used in the development of a claimed method for the  
 CC diagnosis of X-linked severe combined immunodeficiency (XSCID),  
 CC in female carriers and male sufferers.  
 SQ Sequence 369 AA;

Query Match 10.6%; Score 188; DB 1; Length 369;  
 Best Local Similarity 26.8%; Pred. No. 5.92e-08;  
 Matches 51; Conservative 52; Mismatches 75; Indels 12; Gaps 11;

Db 59 EVOCFVFNIEYMNCTWSSSPQATNLTLYRYKVSNDNTFQECSHLFSEKITSGCCQIQ 118  
 QY 31 QIQIYYENLETVQVWNAKYSR-TNTHFRFN-GDE-AYDQCTNYLQEGHTSGCLLD 87  
 Db 119 KEDQLQTFVQVLODPKQPORRAVKLNQLNQLVPRAPENLTLSNLSSESOLELRWKS 178  
 QY 88 AEQRDDILXFISIR-NGTH-PVFTASRW-VVYL-KPSPKRVRES-WHQDAVTV--TCSD 140  
 Db 179 IKERCLQYVQYRSNRDRSTELIVNHEPFLSPVDDLKRYFRVRSRY-NPICGSSQ 237  
 QY 141 LSYGDLLEYVQYRSPDTEW-OSKOENTCNVTIEGLDAEKYSEFWVRKAMEDVYGPDTY 199  
 Db 238 WSKWQPVHW 247  
 QY 200 PSDWSEVTCW 209

RESULT 8  
 ID W35294 standard; Protein; 383 AA.  
 AC W35294;  
 DE Murine IL-13 binding chain (first entry)  
 KW Murine IL-13 binding chain of the IL-13 receptor.  
 KW Interleukin-13; IL-13; interleukin-13 receptor binding chain; IL-13bc;  
 KW mediator; IL-13 receptor binding inhibition; IgE-mediated condition;  
 OS Mus sp.  
 FH Key  
 FT Location/Qualifiers  
 FT 1..21  
 FT Peptide  
 FT /label= signal\_sequence  
 FT /note= "putative"  
 FT 22..383  
 FT /label= mature\_protein  
 FT 22..334  
 FT /label= extracellular\_domain  
 FT 335..356  
 FT /label= transmembrane\_domain  
 FT 357..383  
 FT /label= intracellular\_domain  
 PN WO9731946-A1.  
 PD 04-SEP-1997.  
 PF 28-FEB-1997; U03124.  
 PR 01-MAR-1996; US-609572.  
 PA (GEMY ) GENETICS INST INC.  
 PI Collins M, Donaldson D, Fitz L, Neben T, Whitters M,  
 DR Wood C;  
 DR WPI: 97-448632/41.  
 DR N-PSDB: T75213.  
 PT New nucleic acid encoding interleukin-13 receptor binding chain and  
 PT transformed cells - proteins, antibodies and inhibitors, for  
 PT treating immunoglobulin E-mediated diseases, e.g. Graves disease,  
 PT and in diagnosis  
 PS Claim 11; Pages 30-31; 49pp; English.  
 CC The present sequence represents the murine interleukin-13 (IL-13) binding  
 CC chain of the interleukin-13 receptor, designated IL-13bc. IL-13bc acts a  
 CC mediator of the known biological activities of IL-13. Recombinant  
 CC IL-13bc proteins, and antibodies raised against them, are used to  
 CC inhibit the binding of IL-13 to its receptor. They are particularly used  
 CC to treat IgE-mediated conditions, e.g. allergy, asthma and immune complex  
 CC disorders, especially lupus, nephritis, thyroiditis and Grave's disease.  
 CC They are also used to treat immune deficiency (particularly in  
 CC haematopoietic progenitor cells), cancer etc., and to increase macrophage  
 CC activation, e.g. in vaccination. To potentiate IL-13 activity, a protein  
 CC with such activity is combined with IL-13bc and the mixture applied,  
 CC in vivo, to a cell expressing at least one chain of the IL-13 receptor  
 CC other than IL-13bc. IL-13bc can also be used in diagnosis to detect  
 CC expression of IL-13, its receptor or binding chain, and to raise specific  
 CC antibodies which may be useful for treating some tumours.  
 SQ Sequence 383 AA;

Query Match 7.4%; Score 132; DB 1; Length 383;  
 Best Local Similarity 30.0%; Pred. No. 8.30e-03;  
 Matches 27; Conservative 25; Mismatches 31; Indels 7; Gaps 6;

Db 123 GISDEGSELETKIDMKCIYYNWOYLVCSWKPGKTVYSDNTYMTFFVEGLDHAL-QCQDY 181  
 QY 20 ALOGGGAEG-VQ-IQIYYENLETVQVWNAK--YSRTNLTHFRFNG-DEAYDQCTNY 74  
 Db 182 LOHDERNVGCKLSNLDSSDYKDFICVNGS 211  
 QY 75 LLOEGHTSGCLLDAEQRDDIL-YFSIRNGT 103

RESULT 9  
 ID R78613 standard; Protein; 576 AA.  
 AC R78613;  
 DE 20-FEB-1996 (first entry)  
 KW Expression vector pME18S/mFas-EXT-AIC2A protein prod.  
 KW Expression vector; pME18S/mFas-EXT-AIC2A; murine Fas antigen;

extracellular; region; AIC2A; soluble membrane protein; antibody production; diseases; treatment; prevention.

OS Mus musculus.

PN J07115988-A.

PD 09-MAY-1995.

PF 26-OCT-1993; 267644.

PR 26-OCT-1993; JP-267644.

PA (NIBS ) JAPAN TOBACCO INC.

DR WPI: 95-202847/27.

DR N-PSDB: Q95306.

PT Preparation of soluble membrane proteins - for their use in antibody production for the treatment and prevention of related diseases

PS Claim 10; Pages 36-38; 51pp; Japanese.

CC R78613 is the protein prod. of the expression vector pME18S/murine Fas antigen.extracellular region-AIC2A. The expression vector was used for the prodn. of recombinant soluble membrane proteins. The proteins can be used in antibody prodn. for the treatment and prevention of related diseases.

CC Diseases.

CC Sequence 576 AA;

Query Match 6.5%; Score 115; DB 1; Length 576;

Best Local Similarity 32.4%; Pred. No. 2.45e-01;

Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 502 YIDHTFOVQYKKKSES-WKDSKTENLGRVNSMDLPQLEPDTSYCARVRVKPISD-YD-GI 558

QY 143 YGDLLEYQVYRSPFDTEW-QSKQENTCNV-TIE-G-LDAEKCYSFVVRVKAMEDVYGPDT 198

Db 559 W-SENSNEYTW 568

QY 199 YPSDWSEVTCW 209

RESULT 10

ID R78616 standard; Protein; 596 AA.

AC R78616;

DC 20-FEB-1996 (first entry)

DE Expression vector pME18S/mpd-1.EXT-AIC2A protein prod.

KW Expression vector; pME18S/mpd-1.EXT-AIC2A;

KW extracellular; region; AIC2A; soluble membrane protein;

KW antibody production; diseases; treatment; prevention.

OS Mus musculus.

PN J07115988-A.

PD 09-MAY-1995.

PF 26-OCT-1993; 267644.

PR 26-OCT-1993; JP-267644.

PA (NIBS ) JAPAN TOBACCO INC.

DR WPI: 95-202847/27.

DR N-PSDB: Q95310.

PT Preparation of soluble membrane proteins - for their use in antibody production for the treatment and prevention of related diseases

PS Claim 10; Pages 44-46; 51pp; Japanese.

CC R78616 is the protein prod. of the expression vector pME18S/mpd-1 .extracellular region-AIC2A. The expression vector was used for the prodn. of recombinant soluble membrane proteins. The proteins can be used in antibody prodn. for the treatment and prevention of related diseases.

CC Diseases.

CC Sequence 596 AA;

Query Match 6.5%; Score 115; DB 1; Length 596;

Best Local Similarity 32.4%; Pred. No. 2.45e-01;

Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 517 YIDHTFOVQYKKKSES-WKDSKTENLGRVNSMDLPQLEPDTSYCARVRVKPISD-YD-GI 573

QY 143 YGDLLEYQVYRSPFDTEW-QSKQENTCNV-TIE-G-LDAEKCYSFVVRVKAMEDVYGPDT 198

Db 574 W-SENSNEYTW 583

QY 199 YPSDWSEVTCW 209

RESULT 11

ID R78616 standard; Protein; 600 AA.

AC R78616;

DC 08-SEP-1996 (first entry)

DE Fas antigen #1.

DE Fas antigen; immunocassay; monoclonal antibody; autoimmune disease; SLE; rheumatoid arthritis; serum; systemic lupus erythematosus.

KW Synthetic.

OS Key

FT Key

FT peptide

FT 1..16

FT /note= "signal peptide"

FT 17..600

FT /note= "mature Fas antigen #1"

FT protein

PN WO9601277-A1.

PD 18-JAN-1996.

PF 03-MAR-1995; J00349.

PR 06-JUL-1994; JP-154706.

PR 14-FEB-1995; JP-025637.

PA (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.

PA (NIBS ) JAPAN TOBACCO INC.

PI Hachiya T, Noguchi J, Yonehara S;

DR WPI: 96-087635/09.

DR N-PSDB: T16300.

PT Immunocassay method for soluble Fas antigen in body fluids - for diagnosis of autoimmune diseases such as rheumatoid arthritis and systemic lupus erythematosus.

PS Claim 13; Page 73-77; 124pp; Japanese.

CC R95226 and R95227 represent soluble Fas antigens. These soluble Fas antigen is included in the immunoassay kit of the invention. The kit is for the assay of soluble Fas antigen and contains an immobilised anti-soluble Fas monoclonal antibody, as well as one of these standard soluble Fas antigens. The assay is simple and has high accuracy, high sensitivity, and is capable of assaying a number of different samples (such as serum) and is useful for diagnosis of autoimmune diseases such as rheumatoid arthritis or systemic lupus erythematosus (SLE).

CC Sequence 600 AA;

Query Match 6.5%; Score 115; DB 1; Length 600;

Best Local Similarity 32.4%; Pred. No. 2.45e-01;

Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 521 YIDHTFOVQYKKKSES-WKDSKTENLGRVNSMDLPQLEPDTSYCARVRVKPISD-YD-GI 577

QY 143 YGDLLEYQVYRSPFDTEW-QSKQENTCNV-TIE-G-LDAEKCYSFVVRVKAMEDVYGPDT 198

Db 578 W-SENSNEYTW 587

QY 199 YPSDWSEVTCW 209

RESULT 12

ID R78610 standard; Protein; 600 AA.

AC R78610;

DC 19-FEB-1996 (first entry)

DE Expression vector pME18S/hFas.EXT-AIC2A protein prod.

KW Expression vector; pME18S/hFas.EXT-AIC2A; human Fas antigen; extracellular; region; AIC2A; soluble membrane protein;

KW antibody production; diseases; treatment; prevention.

OS Homo sapiens.

FT Key

FT peptide

FT 1..16

FT /label= sig\_peptide

FT 17..600

FT /label= mat\_peptide

FT J07115988-A.

PD 09-MAY-1995.

PF 26-OCT-1993; 267644.

PR 26-OCT-1993; JP-267644.

PA (NIBS ) JAPAN TOBACCO INC.

DR WPI: 95-202847/27.

DR N-PSDB: Q95301.

PT Preparation of soluble membrane proteins - for their use in antibody production for the treatment and prevention of related diseases

New method of modifying polypeptide hormone-receptor complex - to produce human growth hormone variant, useful for stimulating lactogenic and somatogenic response  
Disclosure; Page 41; 74pp; English.  
This truncated human growth hormone binding protein was used as the basis of mutagenesis experiments to generate receptor variants. The results obtained with the shorter receptor were indistinguishable from those obtained with the 246 amino acid hGHbp. The sequence was modified at the metal binding site. Modification of metal ion chelation allows the response of a cell, organ or animal to a hormone to be controlled, e.g. Zn ion binding is required for tight binding of hGH to the prolactin receptor but not to the hGH receptor. See e.g. R24274 for an example of a preferred variant.  
Sequence 237 AA.

Thu May 11 06:49:29 2000

Query Match 6.3%; Score 111; DB 1; Length 237;  
Best Local Similarity 33.3%; Pred. No. 5.34e-01;  
Matches 17; Conservative 11; Mismatches 22; Indels 1; Gaps 1;  
Db 172 LEYELQYKEVNETKWKMDPILTTSPVYSLKVDKEYEVRVRSKORNSGYG 222  
QY 146 LLYEVQYRSPDTEWQSKQEN-TCNVITIEGLDAEKCYSFVVRKAMEDVYG 195

Search completed: Wed May 10 11:26:55 2000  
Job time : 13 secs.





\*\*\*\*\*  
Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd  
\*\*\*\*\*  
MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:27:14 2000; MasPar time 49.22 Seconds  
Tabular output not generated.  
Title: >US-09-376-430-2  
Description: (1-231) from US09376430A.pap (3 of 25)  
Perfect Score: 1773  
Sequence: 1 MGRVLVWGAAVFLGGWMA.....GEIRDACAETFTPPKPKLSK 231  
Scoring table: PAM 150  
Gap 11  
Searched: 131253 seqs, 12956647 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: a-issued  
1:5A\_COMB 2:5B\_COMB 3:PCT\_COMB 4:backfiles1  
Statistics: Mean 30.786; Variance 128.556; scale 0.239  
pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.  
SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	195	11.0	230	1	US-08-595-Sequence 11, Applicati	2,35e-08
2	195	11.0	230	1	US-08-595-Sequence 11, Applicati	2,35e-08
3	195	11.0	252	1	US-08-595-Sequence 9, Applicati	2,35e-08
4	195	11.0	252	1	US-08-595-Sequence 9, Applicati	2,35e-08
5	195	11.0	347	1	US-08-595-Sequence 7, Applicati	2,35e-08
6	195	11.0	347	1	US-08-595-Sequence 7, Applicati	2,35e-08
7	195	11.0	369	1	US-08-595-Sequence 4, Applicati	2,35e-08
8	195	11.0	369	1	US-08-595-Sequence 4, Applicati	2,35e-08
9	188	10.6	369	3	PCT-US94-0Sequence 69, Applicati	1,00e-07
10	188	10.6	369	2	US-08-424-Sequence 2, Applicati	7,49e-03
11	132	7.4	383	1	US-07-960-Sequence 2, Applicati	2,74e-01
12	113	6.4	897	1	US-08-609-Sequence 4, Applicati	3,51e+00
13	102	5.8	980	2	US-08-673-Sequence 3, Applicati	2,40e+01
14	99	5.6	1239	1	US-08-026-Sequence 4, Applicati	2,02e+01
15	90	5.1	325	2	US-09-003-Sequence 2, Applicati	2,02e+01
16	88	5.0	446	2	US-08-485-Sequence 16, Applicati	2,40e+01
17	89	5.0	446	2	US-08-485-Sequence 16, Applicati	2,40e+01
18	88	5.0	983	1	US-08-162-Sequence 5, Applicati	2,85e+01
19	88	5.0	983	1	US-08-162-Sequence 5, Applicati	2,85e+01
20	87	4.9	488	2	US-08-599-Sequence 4, Applicati	2,85e+01
21	87	4.9	558	2	US-08-825-Sequence 2, Applicati	2,85e+01
22	87	4.9	708	2	US-09-038-Sequence 2, Applicati	2,85e+01
23	87	4.9	708	2	US-09-038-Sequence 2, Applicati	2,85e+01

24	87	4.9	708	3	PCT-US95-0Sequence 2, Applicatio	2,85e+01
25	87	4.9	708	1	US-07-797-Sequence 2, Applicatio	2,85e+01
26	87	4.9	708	1	US-08-308-Sequence 2, Applicatio	2,85e+01
27	87	4.9	918	1	US-08-825-Sequence 6, Applicatio	4,00e+01
28	85	4.8	288	1	US-08-375-Sequence 9, Applicatio	4,00e+01
29	85	4.8	288	1	US-08-752-Sequence 9, Applicatio	4,00e+01
30	85	4.8	1214	2	US-08-486-Sequence 54, Applicati	4,00e+01
31	85	4.8	1214	2	US-08-231-Sequence 54, Applicati	4,00e+01
32	85	4.8	1219	2	US-08-486-Sequence 50, Applicati	4,00e+01
33	85	4.8	1219	2	US-08-231-Sequence 50, Applicati	4,00e+01
34	85	4.8	1231	2	US-08-231-Sequence 48, Applicati	4,00e+01
35	85	4.8	1231	2	US-08-486-Sequence 6, Applicati	4,00e+01
36	85	4.8	1236	2	US-08-486-Sequence 6, Applicati	4,00e+01
37	85	4.8	1236	2	US-08-231-Sequence 52, Applicati	4,00e+01
38	85	4.8	1239	2	US-08-486-Sequence 52, Applicati	4,00e+01
39	85	4.8	1239	2	US-08-231-Sequence 46, Applicati	4,00e+01
40	85	4.8	1244	2	US-08-486-Sequence 46, Applicati	4,00e+01
41	85	4.8	1244	2	US-08-231-Sequence 33, Applicati	4,00e+01
42	83	4.7	187	1	US-08-312-Sequence 33, Applicati	5,60e+01
43	83	4.7	187	3	PCT-US94-0Sequence 2, Applicatio	5,60e+01
44	83	4.7	574	1	US-08-312-Sequence 2, Applicatio	5,60e+01
45	83	4.7	1447	3	PCT-US94-0Sequence 2, Applicatio	5,60e+01

ALIGNMENTS

Sequence 11, Application US/08595974

Sequence 11, Application US/08595974

Patent No. 5705608

GENERAL INFORMATION:

APPLICANT: SUGAMURA, KAZUO

APPLICANT: TAKESHITA, TOSHIKAZU

APPLICANT: ASAO, HIRONOBU

APPLICANT: NAKAMURA, MASATKA

APPLICANT: SHIMAMURA, TOSHIRO

APPLICANT: SUZUKI, MANABU

APPLICANT: HAMURO, JUNJI

TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,

ADDRESSER: P.C.

STREET: 1755 S. Jefferson Davis Highway, Suite 400

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/595,974

FILING DATE: 06-FEB-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/052,205

FILING DATE: 22-APR-1993

APPLICATION NUMBER: JP 104947/1992

FILING DATE: 23-APR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5705608man F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 10-615-0X



Page 3

```

CC ZIP: 22202
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/052,205
CC FILING DATE: 19930422
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: JP 104947/1992
CC FILING DATE: 23-APR-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Oblon, No. 5510259man F.
CC REGISTRATION NUMBER: 24,618
CC REFERENCE/DOCKET NUMBER: 10-615-OX
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703) 413-3000
CC TELEFAX: (703) 413-2220
CC TELEX: 248855 OPAT UR
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 252 amino acids
CC TYPE: AMINO ACID
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 252 AA; 25505 MW; 347903 CN;

Query Match 11.08; Score 195; DB 1; Length 252;
Best Local Similarity 25.98; Pred No. 2,35e-08;
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;

Db 59 EVQCFVNEVMCTNNSSEPOPTNLTHWYKSDNKVKCSHYLSEETTSGLQ 118
   :: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||::
Qy 31 QIQIYNLEFVQTWNASKYSR-TNLTFRHF-NGD-EAYDOCTNYLLQEGHSGCLLD 87

Db 119 KKEIHLTYOTFFVQLQDPREPRQATOMLQNVLVPAPENLTLKLSSQELNNRRF 178
   :|||:: |||:: |||:: |||:: |||:: |||:: |||:: |||::
Qy 88 AEQRDDILYSIR-NGTH-PVFATSRMW-VYYL-KPSSPAHVRF-SWHQDAVTVCSDL 142

Db 179 LNHCHLVQYRTDWDHSWTESQVDYRHKFSLSVDGOKRYTFVRSR-FNPICGSAQH 237
   |||||:: |||||:: |||||:: |||||:: |||||:: |||||:: |||||::
Qy 143 YGD-LLAEEYSPFDTEW-QSKOENTCNVTIEGLDAEKCYISWRVKAMEDVIGPDTP 200

Db 238 SEWSHPHW 246
   |::| |
Qy 201 SDWSEVTCW 209

RESULT 5 STANDARD: PRT; 347 AA.
ID US-08-595-974-7
XX xxxxxx
AC
XX
DT
XX
DE Sequence 7, Application US/08595974
XX Sequence 7, Application US/08595974
CC Patent No. 5705608.
CC GENERAL INFORMATION:
CC APPLICANT: SUGAMURA, KAZUO
CC APPLICANT: TAKESHITA, TOSHIKAZU
CC APPLICANT: ASAO, HIRONOBU
CC APPLICANT: NAKAMURA, MASATAKA
CC APPLICANT: SHIMAMURA, TOSHIO
CC APPLICANT: SUZUKI, MANABU
CC APPLICANT: HAMURO, JUNJI
CC TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE
CC NUMBER OF SEQUENCES: 21
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

```

CC ADDRESSEE: P.C.  
 CC STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 CC CITY: Arlington  
 CC STATE: Virginia  
 CC COUNTRY: U.S.A.  
 CC ZIP: 22202  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent In Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/595,974  
 CC FILING DATE: 06-FEB-1996  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/052,205  
 CC FILING DATE: 22-APR-1993  
 CC APPLICATION NUMBER: JP 104947/1992  
 CC FILING DATE: 23-APR-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Obolon, No. 5705608man F.  
 CC REGISTRATION NUMBER: 24,618  
 CC REFERENCE/DOCKET NUMBER: 10-615-0X  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (703) 413-3000  
 CC TELEFAX: (703) 413-2220  
 CC TELEX: 248855 OPAT UR  
 CC INFORMATION FOR SEQ ID NO: 7:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 347 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 347 AA; 39920 MW; 666553 CN;

Query Match 11.0%; Score 195; DB 1; Length 347;  
 Best Local Similarity 25.9%; Pred. No. 2.35e-08;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 37 EVQCFVFNVEYMNCTWSSSEPOPTNLTHVYKNSDNDKVKQKSHYLFSEITSGCOLQ 96  
 QY 31 QIQIIFYNLEIVQVWTKASKYR-TNLTFFHYRF-NGD-EAYDQCTNYLLQEGHTSGCLLD 87  
 Db 97 KKEIHYQTFVVLQDPPRRQATQMLKQLNVLIPWAPENLTLLHKLSEQLNWNRRF 156  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYLL-KPSPKHVRF-SWHQDAVTVTCSDL 142  
 Db 157 LNHCHLHVQYRTDWDHSHWTEQSVDRHKLPSVDGQKRYTFVRSR-FNPLCGSAQHW 215  
 QY 143 YGD-LLYEVQYRSPDTEW-OSKQENTCNVTIEGLDAEKCYSFWRVKAMEDVYGPDTYP 200  
 Db 216 SEWSHPH 224  
 QY 201 SDWSEVTCW 209

RESULT 6  
 ID US-08-052-205-7 STANDARD; PRT; 347 AA.  
 XX  
 AC xxxxxx  
 XX  
 DT  
 XX

Sequence 7, Application US/08052205  
 Sequence 7, Application US/08052205  
 Patent No. 5510259  
 GENERAL INFORMATION:  
 CC APPLICANT: SUGAMURA, KAZUO  
 CC APPLICANT: TAKESHITA, TOSHIKAZU  
 CC APPLICANT: ASAO, HIRONOBU  
 CC APPLICANT: NAKAMURA, MASATAKA

CC CC  
 CC APPLICANT: SHIMAMURA, TOSHIRO  
 CC APPLICANT: SUZUKI, MANABU  
 CC TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE  
 CC NUMBER OF SEQUENCES: 21  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,  
 CC ADDRESSEE: P.C.  
 CC STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 CC CITY: Arlington  
 CC STATE: Virginia  
 CC COUNTRY: U.S.A.  
 CC ZIP: 22202  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent In Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/052,205  
 CC FILING DATE: 19930422  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: JP 104947/1992  
 CC FILING DATE: 23-APR-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Obolon, No. 5510259man F.  
 CC REGISTRATION NUMBER: 24,618  
 CC REFERENCE/DOCKET NUMBER: 10-615-0X  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (703) 413-3000  
 CC TELEFAX: (703) 413-2220  
 CC TELEX: 248855 OPAT UR  
 CC INFORMATION FOR SEQ ID NO: 7:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 347 amino acids  
 CC TYPE: AMINO ACID  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 347 AA; 39920 MW; 666553 CN;

Query Match 11.0%; Score 195; DB 1; Length 347;  
 Best Local Similarity 25.9%; Pred. No. 2.35e-08;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 37 EVQCFVFNVEYMNCTWSSSEPOPTNLTHVYKNSDNDKVKQKSHYLFSEITSGCOLQ 96  
 QY 31 QIQIIFYNLEIVQVWTKASKYR-TNLTFFHYRF-NGD-EAYDQCTNYLLQEGHTSGCLLD 87  
 Db 97 KKEIHYQTFVVLQDPPRRQATQMLKQLNVLIPWAPENLTLLHKLSEQLNWNRRF 156  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYLL-KPSPKHVRF-SWHQDAVTVTCSDL 142  
 Db 157 LNHCHLHVQYRTDWDHSHWTEQSVDRHKLPSVDGQKRYTFVRSR-FNPLCGSAQHW 215  
 QY 143 YGD-LLYEVQYRSPDTEW-OSKQENTCNVTIEGLDAEKCYSFWRVKAMEDVYGPDTYP 200  
 Db 216 SEWSHPH 224  
 QY 201 SDWSEVTCW 209

RESULT 7  
 ID US-08-595-974-4 STANDARD; PRT; 369 AA.  
 XX  
 AC xxxxxx  
 XX  
 DT  
 XX

Sequence 4, Application US/08595974  
 Sequence 4, Application US/08595974  
 Patent No. 5705608





US-09-376-430-2-03.ra1

Thu May 11 06:49:29 2000

Db 119 KEDIQYOTFFVOLQDPOKQORAVOKLNLONLVIAPENLTLNLSELSOLELRWKSRR 178  
 QY 88 AEQDDILYFIR-NGTH-PVFTASRM-VYIL-KPSPKHVRS-WHODAVIV--TCSD 140  
 Db 179 IKERCLOYLVQYRSNRDRSTELIVNEHPRFSLPSDELKRYTFVRVRSY-NPICGSSQ 237  
 QY 141 LSYGDLLEYQYRSPFTEW-QSKQENTCNVTIEGLDAEKCYSEFWVRKAMEDVYGPTY 199  
 Db 238 WSKWSQPVHW 247  
 QY 200 PSDWSEVTCW 209

RESULT 11  
 ID US-08-609-572-2 STANDARD; PRT; 383 AA.  
 XX  
 AC  
 XX  
 XX  
 DT  
 XX  
 DE  
 XX  
 XX

Sequence 2, Application US/08609572  
 Patent No. 5710023  
 GENERAL INFORMATION:

APPLICANT: Collins, Mary  
 APPLICANT: Donaldson, Debra  
 APPLICANT: Filtz, Lori  
 APPLICANT: Neben, Tamlyn  
 APPLICANT: Whitters, Matthew  
 APPLICANT: Wood, Clive  
 TITLE OF INVENTION: CYTOKINE RECEPTOR CHAIN  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Genetics Institute, Inc.  
 STREET: 87 CambridgePark Drive  
 CITY: Cambridge  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02140  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/609,572  
 FILING DATE:  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brown, Scott A. 32,724  
 REGISTRATION NUMBER: G15268  
 REFERENCE/DOCKET NUMBER:  
 TELEPHONE: (617) 498-8224  
 TELEFAX: (617) 876-5851  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 383 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE 383 AA; 44482 MW; 795803 CN;

Query Match 7.4%; Score 132; DB 1; Length 383;  
 Best Local Similarity 30.0%; Pred. No. 7.49e-03;  
 Matches 27; Conservative 25; Mismatches 31; Indels 7; Gaps 6;  
 Db 123 GISDEGSLKIDGMKCIYNNWOLVCSWPKGTIVSDTYMTFFWVEGLDHAL-OCADY 181  
 QY 20 ALGGGAAEG-VQ-IQIIYENLTVQVWNASK--YSRTNLTFHYRENG-DEAYDQCTNY 74

Db 182 LOHDEKNVCKLSNLDSSDYKDFFCVNGS 211  
 QY 75 LLOEGHTSGCLLDAEQRDIL-YFSIRNGT 103  
 RESULT 12  
 ID US-07-960-389-2 STANDARD; PRT; 897 AA.  
 XX  
 AC  
 XX  
 XX  
 DT  
 XX  
 DE  
 XX  
 XX

Sequence 2, Application US/07960389  
 Patent No. 5705611  
 GENERAL INFORMATION:  
 APPLICANT: HAYASHIDA, Kasuhiro;  
 TITLE OF INVENTION: Human GM-CSF Receptor Component  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Schering-Plough Corporation  
 STREET: 2000 Galloping Hill Road  
 CITY: Kenilworth  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07033  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy Disc  
 COMPUTER: Apple Macintosh  
 OPERATING SYSTEM: System Software 7.1  
 SOFTWARE: Microsoft Word 5.1a  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/960,389  
 FILING DATE: 07-JAN-1993  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 554,745  
 FILING DATE: 18-JUL-1990  
 APPLICATION NUMBER: PCT/US 91/04846  
 FILING DATE: 16-JUL-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Blasdale, John H. C.  
 REGISTRATION NUMBER: 31,895  
 REFERENCE/DOCKET NUMBER: DX0143Q  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (908) 298-2902  
 TELEFAX: (908) 298-5388  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 897 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 FEATURE:  
 OTHER INFORMATION: Human GM-CSF receptor; Signal Sequence: -17 to -1;  
 SEQUENCE 897 AA; 97323 MW; 4443497 CN;

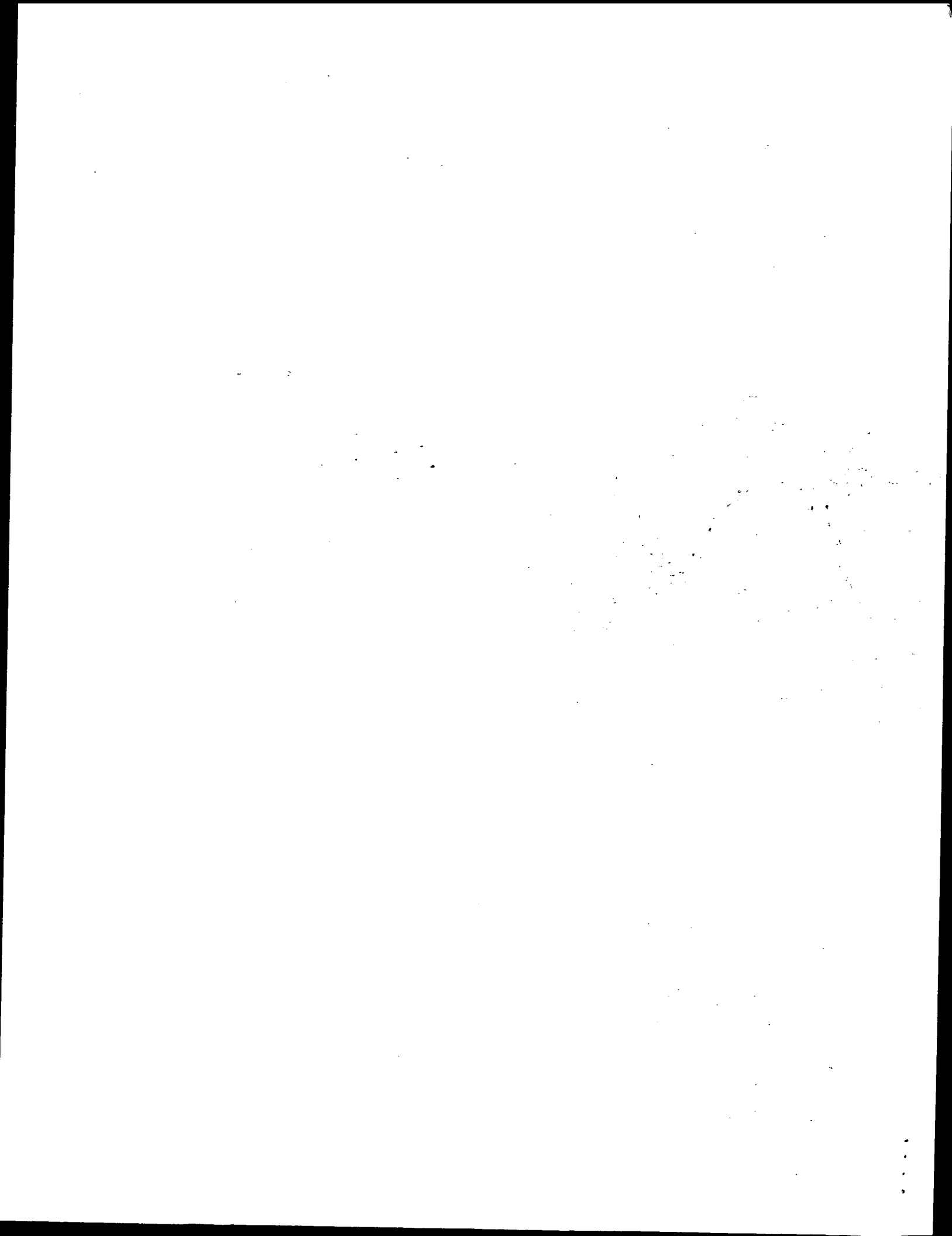
Query Match 6.4%; Score 113; DB 1; Length 897;  
 Best Local Similarity 31.9%; Pred. No. 2.74e-01;  
 Matches 22; Conservative 10; Mismatches 34; Indels 3; Gaps 3;  
 Db 372 FEIQYRKDRAT-WKDSKTTETLQNAHSMALPEPSTRYWARVVRTSRTGYNGIWESEW 430  
 QY 148 YEVQYKSPFDEW-QSKQENTCNVTIEGLDA-EKCYSEFWVRKAMEDVGPDIYPSDWE 205  
 Db 431 ARSWDTESV 439  
 QY 206 VTCWQGEI 214  
 RESULT 13





Thu May 11 06:49:29 2000

CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Nishiohata Residence 1-107  
CC STREET: 5214, Nishiohata-machi  
CC CITY: Niigata-shi  
CC STATE: Niigata-ken  
CC COUNTRY: JAPAN  
CC ZIP: 951  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
CC COMPUTER: IBM Compatible  
CC OPERATING SYSTEM: MS-DOS v.5  
CC SOFTWARE: Word Perfect 5.1  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/026,138E  
CC FILING DATE: 26-FEB-1993  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: JP 39563/1992  
CC FILING DATE: 26-FEB-1992  
CC APPLICATION NUMBER: JP 173155/1992  
CC FILING DATE: 30-JUN-1992  
CC APPLICATION NUMBER: JP 215017/1992  
CC FILING DATE: 12-AUG-1992  
CC APPLICATION NUMBER: JP 303878/1992  
CC FILING DATE: 13-NOV-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Hamburg, C.Bruce  
CC REGISTRATION NUMBER: 22,389  
CC REFERENCE/DOCKET NUMBER: F-4551  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 986-2340  
CC TELEFAX: (212) 953-7733  
CC INFORMATION FOR SEQ ID NO: 3:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1239 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single strand  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC ORIGINAL SOURCE:  
CC ORGANISM: mouse  
CC TISSUE TYPE: brain  
CC PUBLICATION INFORMATION:  
CC AUTHORS: Masayoshi MISHINA  
CC TITLE: NOVEL PROTEINS AND GENES CODING THE SAME  
CC RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 to 1239  
CC SQ SEQUENCE 1239 AA; 135308 MW; 8004364 CN;  
  
CC Query Match 5.1%; Score 90; DB 1; Length 1239;  
CC Best Local Similarity 34.8%; Pred. No. 1.70e+01;  
CC Matches 8; Conservative 9; Mismatches 6; Indels 0; Gaps 0;  
  
CC Db 14 LLGAWAGLGAGGCGEQAIVTAVVF 36  
CC QY 14 LLGGWALGGGGAEGVQIQIYY 36  
  
CC Search completed: Wed May 10 11:28:11 2000  
CC JOB time : 57 secs.



\*\*\*\*\*  
M P S R L H  
\*\*\*\*\*  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:26:03 2000; MasPar time 16.65 Seconds  
Tabular output not generated. 654.425 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (1-231) from US09376430A.pep (3 of 25)  
Perfect score: 1773  
Sequence: 1 MGRLLVLLWGAFLGLGWA.....GEIRDCAETPTPKPLSK 231

Scoring table: PAM 150  
Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: p1r62  
1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 44.063; Variance 78.850; scale 0.559

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	210	11.8	373	2	A55718 interleukin-2 recepto	7.96e-21
2	195	11.0	369	2	A42565 interleukin-2 recepto	6.16e-18
3	188	10.6	369	2	I49280 interleukin-2 recepto	1.31e-16
4	115	6.5	878	1	A40091 interleukin-3 recepto	7.51e-04
5	113	6.4	897	1	A39255 cytokine receptor com	1.53e-03
6	109	6.1	435	2	probable aspartate am	6.24e-03
7	108	6.1	435	2	E71350 cytokine receptor com	8.82e-03
8	104	5.9	430	2	A35782 cytokine receptor	3.46e-02
9	100	5.6	638	2	A33991 somatotropin receptor	1.31e-01
10	100	5.6	638	2	S04530 somatotropin receptor	1.31e-01
11	99	5.6	4436	2	E71086 hypothetical protein	1.82e-01
12	99	5.6	6839	2	E57242 twitkin - Caenorhab	1.82e-01
13	97	5.5	302	2	S50579 hypothetical protein	3.49e-01
14	97	5.5	311	2	S77803 hypothetical protein	3.49e-01
15	96	5.5	1409	2	S74916 alkaline phosphatase	2.53e-01
16	96	5.4	290	1	type IV prelipin pept	4.82e-01
17	96	5.4	638	2	S12136 somatotropin receptor	4.82e-01
18	96	5.4	896	2	S156563 somatotropin receptor	4.82e-01
19	94	5.3	285	2	S72220 hypothetical protein	9.09e-01
20	94	5.3	638	2	S28176 somatotropin receptor	9.09e-01
21	94	5.3	1879	2	S74915 extracellular nucleas	9.09e-01
22	94	5.3	26926	1	I38344 titin, cardiac muscle	9.09e-01
23	93	5.2	173	2	T15695 hypothetical protein	1.25e+00

24 92 5.2 349 2 JC6311 interferon receptor-c 1.70e+00  
25 92 5.2 358 2 T07719 aldose 1-epimerase-li 1.70e+00  
26 92 5.2 361 2 JN0716 glutamate--ammonia li 1.70e+00  
27 92 5.2 684 2 S60266 novel antigen recepto 1.70e+00  
28 90 5.1 376 2 D72493 hypothetical protein 3.15e+00  
29 90 5.1 629 2 A30001 ribosomal protein S6 3.15e+00  
30 91 5.1 636 2 H65216 hypothetical 70.9 kD 3.15e+00  
31 90 5.1 1098 2 G70697 probable arabinosyltr 3.15e+00  
32 90 5.1 1203 2 I55466 N-methyl-D-aspartate 3.15e+00  
33 91 1223 2 I38111 phosphorylase kinase 3.32e+00  
34 91 1237 2 A31334 phosphorylase kinase 3.32e+00  
35 90 5.1 1239 2 I49705 glutamate receptor ch 3.32e+00  
36 91 5.1 1241 2 S40528 phosphorylase kinase 3.32e+00  
37 88 5.0 137 2 JQ2325 virion protein Afi1 5.76e+00  
38 89 5.0 255 2 E72226 conserved hypothetical 4.26e+00  
39 88 5.0 273 2 G01418 cytokine receptor fam 5.76e+00  
40 88 5.0 325 2 A47003 cytokine receptor fam 5.76e+00  
41 89 5.0 352 2 G64965 membrane protein yeaA 4.26e+00  
42 88 5.0 562 1 Q0BEH5 phosphotransferase (E 5.76e+00  
43 89 5.0 729 2 T06127 probable sugar transp 4.26e+00  
44 88 5.0 983 2 B45583 receptor tyrosine kin 5.76e+00  
45 89 5.0 1286 2 A42150 P-glycoprotein pgp1 - 4.26e+00

ALIGNMENTS

RESULT 1  
ENTRY #type complete  
TITLE Interleukin-2 receptor gamma chain precursor - dog  
ORGANISM #formal\_name Canis lupus familiaris #common\_name dog  
DATE 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 23-Jul-1999

ACCESSIONS A55718  
REFERENCE A55718  
#authors Henthorn, P.S.; Somberg, R.L.; Fimiani, V.M.; Puck, J.M.; Patterson, D.F.; Felsburg, P.J.  
#journal Genomics (1994) 23:69-74  
#title IL-2Rgamma gene microdeletion demonstrates that canine X-linked severe combined immunodeficiency is a homologue of the human disease.  
#cross-references MUID:95130114  
#accession A55718  
#status preliminary  
#molecule\_type mRNA  
#residues 1-373 #label HEN  
#cross-references GB:U04361; NID:9517411; PIDN:AAC48403.1; PID:9517412  
CLASSIFICATION #superfamily interleukin-2 receptor gamma chain  
KEYWORDS cytokine receptor; duplication  
SUMMARY #length 373 #molecular-weight 42516 #checksum 7960

Query Match 11.8%; Score 210; DB 2; Length 373;  
Best Local Similarity 25.9%; Pred. No. 7.96e-21;  
Matches 49; Conservative 56; Mismatches 73; Indels 11; Gaps 10;

Db 59 EVOCFVFNVEYMMCTWNSSEPRPTNLTTHYWKNSNDKQVCEGHYLFREVTAGCWLQ 118  
QY 31 QIQIIFNLTQVQVWNASKYSR-TNLTTHYRF-NGDEA-YDQCTNYLLOEGHTSGCLD 87  
Db 119 KEIHLHYEFVVLDRPREPROSTOKLQNLVWPAPENLTNLNLSQELSLNSNRH 178  
QY 88 AEORD--DILYFIRNGTHPTFTAS-RWVYVL-KSSPKHVR-F-SWHQDAVITCSDLS 142  
Db 179 LDHCLHVVQYRSDWRSTEQSDVDRNSFLSVQGOKEFYTFVRSRY-NPLCGSAQRW 237  
QY 143 YGD-LLYEVOYRPFPTDEW-QSKQENTCNVTIEGLDAKCYISFWRVKAMEDVIGDPTYP 200  
Db 238 SEWSHPHFW 246  
QY 201 SDWSEVTCW 209

RESULT 2  
ENTRY #type complete

```

interleukin-2 receptor gamma chain - human
#formal_name Homo sapiens common_name man
04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change
23-Jul-1999
ACCESSIONS A42565; A46591; I54332
REFERENCE #authors Takeshita, T.; Asao, H.; Ohtani, K.; Ishii, N.; Kumaki, S.; Tanaka, N.; Munakata, H.; Nakamura, M.; Sugamura, K.
#journal Science (1992) 257:379-382
#title Cloning of the gamma chain of the human IL-2 receptor.
#cross-references MIM:293358
#accession A42565
##status preliminary; not compared with conceptual translation
##molecule_type nucleic acid; protein
##residues 1-369 #label TAK
##cross-references GB:D11086; NID:G303611; PIDN:BAA01857.1;
PID:D1002334; PID:G219890
##experimental_source MOLT beta lymphoid cells
##note sequence extracted from NCBI backbone (NCBIP:109167)
REFERENCE A46591
#authors Noruchi, M.; Adelstein, S.; Cao, X.; Leonard, W.J.
#journal J. Biol. Chem. (1993) 268:13601-13608
#title Characterization of the human interleukin-2 receptor gamma chain gene.
#cross-references MIM:293358
#accession A46591
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 1-369 #label RES
##cross-references GB:L12183; NID:G307056; PIDN:AAA59145.1; PID:G307058
REFERENCE I54332
#authors Puck, J.M.; Deschenes, S.M.; Porter, J.C.; Dutra, A.S.; Brown, C.J.; Willard, H.F.; Henthorn, P.S.
#journal Hum. Mol. Genet. (1993) 2:1099-1104
#title The interleukin-2 receptor gamma chain maps to Xq13.1 and is mutated in X-linked severe combined immunodeficiency, SCIDX1.
#cross-references MIM:293358
#accession I54332
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 1-369 #label RES
##cross-references GB:L19546; NID:G349631; PIDN:AAC37524.1; PID:G349632
GENETICS
#gene GDB:IL2RG; SCIDX1; IMD4
#map_position Xq13.1-Xq13.1
#introns 39/1; 90/2; 152/1; 198/3; 253/1; 285/2; 308/3
#note defects are associated with an X-linked form of severe combined immunodeficiency
#superfamily interleukin-2 receptor gamma chain
#cytokine_receptor duplication; immunodeficiency; severe combined immunodeficiency; transmembrane protein
SUMMARY #length 369 #molecular_weight 42287 #checksum 7850
Query Match 11.0%; Score 195; DB 2; Length 369;
Best Local Similarity 25.9%; Pred. No. 6.16e-18;
Matches 49; Conservativity 52; Mismatches 77; Indels 11; Gaps 11;
Db 59 EVQCFVENVYNCNTWSSEFPQNLTHYKNSDNDKVKCSHYLFSEITSCQLQ 118
Qy 31 QIQIIFNLEIVQVTVNASKYSR-TNLTFHYRF-NGD-EAYDQCTNLLQEGHTSGCLLD 87
Db 119 KKEIHLTYFTVQLODPRPRQATQMLKONLWIPWAPENLTHKLSQLELWNNRF 178
Qy 88 AEQRDILYFSIR-NGTH-PVFTASRW-VYLL-RPSPKPHRF-SWHQDAVTVTCSDL 142
Db 179 LNHCHLHVQHTDWDHSGWTEQSVYRHKFSLPSVDGQKRYTFRRSR-FNPLCGSAQHW 237
Qy 143 YGD-LLYEVQVRSFPDTEW-QSKQENTCNVTTEGLDAEKCYFWRVKAMEDVYGPDTP 200
Db 238 SEWSHPHWH 246

```

---

```

QY 201 SDWSEVTCW 209
RESULT 3
ENTRY I49280 #type complete
TITLE interleukin-2 receptor gamma chain precursor - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 23-Jul-1999
ACCESSIONS I49280; A47514; JN0775; S37582; I53398
REFERENCE A47514
#authors Cao, X.; Kozak, C.A.; Liu, Y.
#journal Proc. Natl. Acad. Sci. U.S.A. (1993) 90:8464-8468
#title Characterization of cDNAs encoding the murine interleukin 2 receptor (IL-2R) gamma chain: Chromosomal mapping and tissue specificity of IL-2R gamma chain expression.
#cross-references MIM:293358
#accession I49280
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 1-369 #label CAO
##cross-references EMBL:D21795; NID:G272349; PIDN:AAA64279.1;
PID:G272350
#accession A47514
##status translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-369 #label KUM
##cross-references GB:L20048; NID:G404067; PIDN:AAA39286.1; PID:G404068
REFERENCE JN0592
#authors Kumaki, S.; Kondo, M.; Takeshita, T.; Asao, H.; Nakamura, M.; Sugamura, K.
#journal Biochem. Biophys. Res. Commun. (1993) 193:356-363
#title Cloning of the mouse interleukin 2 receptor gamma chain: Demonstration of functional differences between the mouse and human receptors.
#cross-references MIM:293358
#accession JN0592
##status nucleic acid sequence not shown
##molecule_type mRNA
##residues 1-369 #label KUM
##cross-references DBJ:D13565; NID:G303684; PIDN:BAA02760.1;
PID:D1003265; PID:G303685
REFERENCE JN0775
#authors Kobayashi, N.; Nakagawa, S.; Minami, Y.; Taniguchi, T.; Kono, T.
#journal Gene (1993) 130:303-304
#title Cloning and sequencing of the cDNA encoding a mouse IL-2 receptor gamma chain.
#cross-references MIM:293358
#accession JN0775
##molecule_type mRNA
##residues 1-369 #label KOB
##cross-references GB:D13821; NID:G436045; PIDN:BAA02974.1;
PID:D1003480; PID:G436046
REFERENCE S37582
#authors Chiu, R.K.; Dougherty, G.J.
#submission submitted to the EMBL Data Library, October 1993
#description Regulation of CD44-mediated cellular adhesion by the IL-2 R gamma chain.
#accession S37582
##status preliminary
##molecule_type mRNA
##residues 1-350, 'S', 352-366, 'S', 368-369 #label CHI
##cross-references EMBL:X75337
REFERENCE I53398
#authors DiSanto, J.P.; Certain, S.; Wilson, A.; MacDonald, H.R.; Avner, P.; Fischer, A.; de Saint Basile, G.
#journal Eur. J. Immunol. (1994) 24:3014-3018
#title The murine interleukin-2 receptor gamma chain gene: organization, chromosomal localization and expression in the adult thymus.
#cross-references MIM:293358
#accession I53398

```

##status preliminary; translated from GB/EMBL/DBDJ  
 ##molecule\_type DNA  
 ##residues 1-369 #label RES  
 ##cross-references GB:S75852; NID:g861554; PIDN:AAB32904.1; PID:g861555  
 GENETICS  
 #gene IL-2Rgamma  
 #introns 39/1; 90/2; 152/1; 199/3; 254/1; 286/2; 308/3  
 COMPLEX The high affinity receptor is a heterotrimer of alpha (see PIR:UHM52), beta (see PIR:A35052), and gamma chains; heterodimers of alpha or beta and gamma chains are intermediate affinity receptors.

FUNCTION  
 #description receptor for interleukin-2  
 #pathway interleukin-2 stimulated growth and differentiation of T cells, B cells, NK cells, LAK cells, monocytes, macrophages, and oligodendrocytes  
 CLASSIFICATION #superfamily interleukin-2 receptor gamma chain  
 KEYWORDS cytokine receptor; duplication; glycoprotein; transmembrane protein

FEATURE  
 1-22 #domain signal sequence #status predicted #label SIG\  
 23-369 #product interleukin-2 receptor gamma chain #status predicted #label MAT\  
 256-284 #domain transmembrane #status predicted #label TM\  
 71, 75, 84, 96, 159, #binding site carbohydrate (Asn) (covalent) #status predicted  
 164, 306 #length 369 #molecular-weight 42241 #checksum 6734  
 SUMMARY

Query Match 10.6%; Score 188; DB 2; Length 369;  
 Best Local Similarity 26.8%; Pred. No. 1.31e-16;  
 Matches 51; Conservative 52; Mismatches 75; Indels 12; Gaps 11;

Db 59 EVQCFVFNENYMCNWSSEPAQNLTLHYRYKYSDNNFTQECSHLYFSKBITSGCQIQ 118  
 QY 31 QIQIIVFNLETQVVTWNSKYSR-TNLTPHYRFN-GDE-AYDQCTNYLQEGHTSGCLLD 87  
 Db 119 KEDIQYQTFVVLQDPQKQRRVQKLNQIVIPRAPENLTLSLSQLELRKSRH 178  
 QY 88 AQQRDILYFSIR-NGTH-PVFTASRW-VYVL-KPSSPKHYRFS-WHDAVIV--TCSD 140  
 Db 179 IKRCLOYLVQVRSNDRSGWTELIVNHEPFLSPVDELRKTRFYRSRY-NPICSSQO 237  
 QY 141 LSGYDLYEVQVRSPPDTEM-OSKQENTCNVTIEGLDAEKCYSFWRVRKAMEDVYGFDTY 199  
 Db 238 WSKWSQPVHW 247  
 QY 200 PSDMSEVTCW 209

RESULT 4  
 ENTRY #type complete  
 TITLE Interleukin-3 receptor beta chain precursor - mouse  
 ORGANISM #formal\_name Mus musculus #common\_name house mouse  
 DATE 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 22-Jun-1999  
 ACCESSIONS A40091; A43022  
 REFERENCE A40091  
 #authors Itoh, N.; Yonehara, S.; Schreurs, J.; Gorman, D.M.; Maruyama K.; Ishii, A.; Yahara, I.; Arai, K.; Miyajima, A.  
 #journal Science (1990) 247:324-327  
 #title Cloning of an interleukin-3 receptor gene: a member of a distinct receptor gene family.

#cross-references MUID:9017145  
 #accession A40091  
 #status nucleic acid sequence not shown  
 #molecule\_type mRNA  
 #residues 1-878 #label ITO  
 REFERENCE A43022  
 #authors Gorman, D.M.  
 #submission submitted to GenBank, November 1989  
 #accession A43022  
 #molecule\_type mRNA

##residues 1-815, 'O', 817-878 #label GOR  
 ##cross-references GB:M29855; NID:g198342; PIDN:AAA39295.1; PID:g309406  
 COMMENT In mice there are two classes of high-affinity IL-3 receptors. One contains this IL-3-specific beta chain and the other contains the beta chain also shared by high-affinity IL-5 and GM-CSF receptors.

CLASSIFICATION #superfamily interleukin-3 receptor beta chain; cytokine receptor homology  
 KEYWORDS cytokine receptor; duplication; transmembrane protein  
 FEATURE 1-22 #domain signal sequence #status predicted #label SIG\  
 23-878 #product interleukin-3 receptor beta chain #status predicted #label MAT\  
 23-440 #domain extracellular #status predicted #label EXT\  
 39-236 #domain cytokine receptor homology #label CRS1\  
 254-433 #domain cytokine receptor homology #label CRS2\  
 441-462 #domain transmembrane #status predicted #label TM\  
 463-878 #domain intracellular #status predicted #label INT  
 SUMMARY #length 878 #molecular-weight 97222 #checksum 2346

Query Match 6.5%; Score 115; DB 1; Length 878;  
 Best Local Similarity 32.4%; Pred. No. 7.51e-04;  
 Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;  
 Db 370 YIDTFQVQYKKSES-WKDSKTENLGRVNSMDLPQLEPDTSYCARVRVKPISD-YD-GI 426  
 QY 143 YGDLLYEVQVRSPPDTEM-OSKQENTCNV-TIE-G-LDAEKCYSFWRVRKAMEDVYGPDT 198  
 Db 427 W-SEWSNEYTW 436  
 QY 199 YPSDWSEVTCW 209

RESULT 5  
 ENTRY #type complete  
 TITLE Cytokine receptor common beta chain precursor - human  
 ORGANISM #formal\_name Homo sapiens #common\_name man  
 DATE 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 02-Sep-1997  
 ACCESSIONS A39255  
 REFERENCE A39255  
 #authors Hayashida, K.; Kitamura, T.; Gorman, D.M.; Arai, K.; Yokota, T.; Miyajima, A.  
 #journal Proc. Natl. Acad. Sci. U.S.A. (1990) 87:9655-9659  
 #title Molecular cloning of a second subunit of the receptor for human granulocyte-macrophage colony-stimulating factor (GM-CSF): reconstruction of a high-affinity GM-CSF receptor.

#cross-references MUID:91088571  
 #accession A39255  
 #molecule\_type mRNA  
 #residues 1-897 #label HAY  
 #cross-references GB:M38275  
 COMMENT The human high-affinity IL-3, IL-5, and GM-CSF receptors have ligand-specific alpha chains and share this common beta chain.

GENETICS  
 #gene GDB:CSF2RB  
 #cross-references GDB:126838; OMIM:138981  
 #map\_position 22q13.1-22q13.1  
 CLASSIFICATION #superfamily interleukin-3 receptor beta chain; cytokine receptor homology  
 KEYWORDS alternative splicing; cytokine receptor; duplication; transmembrane protein

FEATURE  
 1-16 #domain signal sequence #status predicted #label SIG\  
 17-897 #product cytokine receptor common beta chain #status predicted #label MAT\  
 17-443 #domain extracellular #status predicted #label EXT\  
 35-232 #domain cytokine receptor homology #label CRS1\  
 250-431 #domain cytokine receptor homology #label CRS2\  
 444-460 #domain transmembrane #status predicted #label TM\  
 461-897 #domain intracellular #status predicted #label INT  
 SUMMARY #length 897 #molecular-weight 97323 #checksum 753





```

1737-1473,
1568-1670,
1671-1769,
1865-1964,
1965-2065,
2159-2258,
2259-2357,
2451-2550,
2551-2650,
2746-2847,
2848-2947,
3042-3141,
3142-3241,
3337-3435,
3436-3535,
3630-3729,
3730-3828,
3922-4022,
4023-4123,
4215-4313,
4314-4413,
4416-4516,
4612-4710,
4711-4811,
4908-5009,
5010-5109,
5110-5210,
5339-5499,
5500-5600,
5791-5889,
5940-6197,
5948-5956
5971
SUMMARY
#region motif 1\
#domain protein kinase homology #label KIN\
#region protein kinase ATP-binding motif\
#active_site Lys #status predicted
#length 6839 #molecular-weight 753494 #checksum 1785

Query Match 5.6%; Score 99; DB 2: Length 6839;
Best Local Similarity 33.9%; Pred. No. 1.82e-01;
Matches 19; Conservative 8; Mismatches 27; Indels 2; Gaps 2
Db 1707 YIVEVPDPTKEWKVKRVDPDINASIGLKEGKEYQFRVRVKNKAGP-GQPSEPS 1761
QY 148 YEVOYRSPPDTEW-QSKQENTCNVTIEGLDAEKCYSFVVRVKAMEDVYGPDTYPSD 202

RESULT 13
ENTRY #type complete
TITLE hypothetical protein YER076c - yeast (Saccharomyces cerevisiae)
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 21-Nov-1997
ACCESSIONS S50579
REFERENCE S50428
#authors Dietrich, F.S.
#submission submitted to the EMBL Data Library, December 1994
#description The sequence of S. cerevisiae cosmids 9747, 8198, 9781, and lambda clones 3612 and 6052.
#accession S50579
#molecule_type DNA
#residues
#cross-references EMBL:U18839; NID:g603313; MIPS:YER076c
GENETICS
#map_position 5R
SUMMARY #length_302 #molecular-weight 33466 #checksum 9755

Query Match 5.58%; Score 97; DB 2: Length 302;
Best Local Similarity 22.8%; Pred. No. 3.49e-01;
Matches 18; Conservative 21; Mismatches 33; Indels 7; Gaps 7;
Db 194 CGSQEETNIFFDQEGWSLFVKTW-ST-NSSCDITASE-GNLTCA-VRVYSVSMHN-HGK 248
QY 138 CSDLISYGLLYEQVY-RSPFDTEWQSKQENCNVTIEGLDAEKCYSFVVRVKAMEDVYGP 196

```



Thu May 11 06:49:30 2000

Db 249 TAFCVYSHGDSW-RAELR 266  
 QY 197 DTYPDSWSEVTCWGEIR 215

Best Local Similarity 41.7%; Pred. No. 2.53e-01; Mismatches 9; Indels 3; Gaps 3;  
 Matches 15; Conservative

Db 1336 LGNQALAA-AETVQGVN-QVLWQNLDTNQIGVWNS 1369  
 QY 15 LGGWMLGGGAAEGVQIQIIFNFLETQVQV-TWNAS 49

Search completed: Wed May 10 11:26:24 2000  
 Job time : 21 secs.

RESULT 14  
 ENTRY  
 TITLE  
 ORGANISM  
 DATE

S77803 #type complete  
 hypothetical protein MC012 - Mycoplasma capricolum (SGC3)  
 #formal\_name Mycoplasma capricolum  
 09-Oct-1997 #sequence\_revision 31-Oct-1997 #text\_change  
 31-Oct-1997

ACCESSIONS  
 S77803; S49950

REFERENCE  
 S77739  
 Bork, P.; Ouzounis, C.; Casari, G.; Schneider, R.; Sander,  
 C.; Dolan, M.; Gilbert, W.; Gillevet, P.M.  
 Mol. Microbiol. (1995) 16:955-967  
 Exploring the Mycoplasma capricolum genome: a minimal cell  
 reveals its physiology.

#cross-references MUID:96059641  
 #accession S77803  
 #status nucleic acid sequence not shown; translation not shown  
 #molecule\_type DNA  
 #residues 1-311 #label BOR  
 #cross-references EMBL:233015; NID:g599860; PID:g602031  
 #experimental\_source ATCC 27343  
 #note the nucleotide sequence was submitted to the EMBL Data  
 Library, July 1994

GENETICS  
 #genetic\_code SGC3  
 #length 311 #molecular\_weight 35768 #checksum 4450

Query Match 5.5%; Score 97; DB 2; Length 311;  
 Best Local Similarity 27.9%; Pred. No. 3.49e-01;  
 Matches 12; Conservative 11; Mismatches 19; Indels 1; Gaps 1;

Db 202 FCLEVVKAIREVIDKAPKNIFGFRAPTEETYGDLGYTIED 244  
 QY 37 FNLETVQVTWNA-SKYSRTNLTFFRNGDEAYDQCTNYLQE 78

RESULT 15  
 ENTRY  
 TITLE  
 ALTERNATE\_NAMES  
 ORGANISM  
 #variety  
 DATE

S74916 #type complete  
 alkaline phosphatase - Synechocystis sp. (strain PCC 6803)  
 protein sl10654  
 #formal\_name Synechocystis sp.  
 PCC 6803  
 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change  
 21-Aug-1998

ACCESSIONS  
 S74916  
 S74322

REFERENCE  
 Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.;  
 Nakamura, Y.; Miyajima, N.; Hirose, M.; Sugita, M.;  
 Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.;  
 Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimo,  
 S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.;  
 Yasuda, M.; Tabata, S.

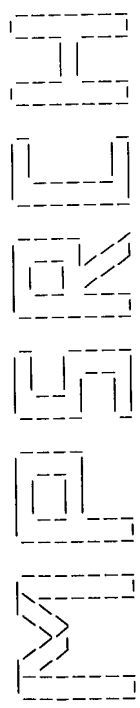
#journal DNA Res. (1996) 3:109-136  
 #title Sequence analysis of the genome of the unicellular  
 cyanobacterium Synechocystis sp. PCC6803. II. Sequence  
 determination of the entire genome and assignment of  
 potential protein-coding regions.  
 #cross-references MUID:97061201

#accession S74916  
 #status preliminary  
 #molecule\_type DNA  
 #residues 1-1409 #label KAN  
 #cross-references EMBL:D90902; GB:AB001339; NID:gl652027; PID:dl017689  
 #note the nucleotide sequence was submitted to the EMBL Data  
 Library, June 1996

SUMMARY  
 #length 1409 #molecular\_weight 149389 #checksum 9390

Query Match 5.5%; Score 98; DB 2; Length 1409;





(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:19:44 2000; Maspar time 76.56 Seconds  
Tabular output not generated. 91.891 Million cell updates/sec

Title: >US-09-376-430-2  
(1-231) from US09376430A.pap (3 of 25)  
Perfect Score: 1773  
Sequence: 1 MGRVLVLLWGAAYVLLGGWMA.....GEIRDCAETPTPPKPLSK 231

Scoring table: PAM 150  
Gap 11  
Searched: 83857 seqs, 30454973 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: swiss-prot38  
1:swissprot

Statistics: Mean 44.839; Variance 72.373; scale 0.620

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	210	11.8	373	1	CYRG_CANFA	2.62e-27
2	195	11.0	369	1	CYRG_HUMAN	3.99e-20
3	188	10.6	369	1	CYRG_MOUSE	1.16e-18
4	187	10.5	379	1	CYRG_BOVIN	1.17e-18
5	115	6.5	878	1	IL13_MOUSE	1.19e-04
6	113	6.4	897	1	CYRB_HUMAN	2.60e-04
7	108	6.1	896	1	CYRB_MOUSE	1.76e-03
8	102	5.8	380	1	IL132_HUMAN	1.63e-02
9	100	5.6	638	1	GHR_HUMAN	3.35e-02
10	97	5.5	302	1	YEP6_YEAST	9.72e-02
11	96	5.4	290	1	LEP3_AERHY	1.38e-01
12	96	5.4	638	1	GHR_PIG	2.75e-01
13	94	5.2	361	1	GLNA_PANAR	5.43e-01
14	92	5.2	722	1	Y390_HAELN	3.87e-01
15	93	5.2	427	1	INTERLEUKIN-13 RECEPTOR	1.06e-00
16	90	5.1	629	1	RIBOSOMAL PROTEIN S6 K	1.06e-00
17	90	5.1	629	1	Y390_HAELN	7.60e-01
18	91	5.1	1223	1	PHOSPHORYLASE B KINASE	7.60e-01
19	91	5.1	1237	1	KPBL_HUMAN	1.06e-01
20	91	5.1	1237	1	KPBL_RABIT	1.06e-01
21	90	5.1	1239	1	NME3_MOUSE	7.60e-01
22	91	5.1	1241	1	KPBL_MOUSE	7.60e-01
23	88	5.0	325	1	CYRG_HUMAN	2.04e-00

ID	CYRG_CANFA	STANDARD	PRT	373 AA
AC	P40321	01-PEB-1995 (Rel. 31, Created)		
DT	01-PEB-1995 (Rel. 31, Last sequence update)			
DT	01-PEB-1995 (Rel. 31, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C)			
DE	(INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64).			
GN	IL2RG, familiaris (Dog).			
OS	Canis familiaris (Dog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-SPLEEN;			
RX	MEDLINE: 95130114.			
RA	Henthorn P.S., Somberg R.L., Fimiani V.M., Puck J.M., Patterson D.F.,			
RA	Felsburg P.J.;			
RT	"IL-2R gamma gene microdeletion demonstrates that canine X-linked			
RT	severe combined immunodeficiency is a homologue of the human			
RT	disease."			
RL	Genomics 23:69-74(1994).			
CC	FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF			
CC	INTERLEUKINS.			
CC	-1- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND			
CC	PROBABLY ALSO THE IL-13 RECEPTORS.			
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-1- DISEASE: DEFECTS IN IL2RG ARE THE CAUSE OF A CANINE X-LINKED			
CC	SEVERE COMBINED IMMUNODEFICIENCY.			
CC	-1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.			
CC	-1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; U04361; AAC48403.1;			
DR	HSP; P31785; IILN.			
DR	PROSITE; PS00241; RECEPTOR_CYTOKINES_1; 1.			
DR	PROSITE; PS00340; RECEPTOR_CYTOKINES_2; FALSE_NEG.			
DR	PFAM; PF00041; fn3; 1.			
KW	Receptor; Transmembrane; Glycoprotein; signal.			
FT	POTENTIAL.			

RP IDENTIFICATION AS A IL-4R SUBUNIT.  
RX MEDLINE; 94090315.  
RA Kondo M., Takeshita T., Ishii N., Nakamura M., Watanabe S.,  
RA Arai K.-I., Sugamura K.;  
RT "Sharing of the interleukin-2 (IL-2) receptor gamma chain between  
RT receptors for IL-2 and IL-4.";  
RL Science 262:1874-1877(1993).  
RN [5]  
RP IDENTIFICATION AS A IL-4R SUBUNIT.  
RX MEDLINE; 94090317.  
RA Russell S.M., Kegan A.D., Harada N., Nakamura Y., Noguchi M.,  
RA Leland P., Friedmann M.C., Miyajima A., Puri R.K., Paul W.E.,  
RA Leonard W.J.;  
RT "Interleukin-2 receptor gamma chain: a functional component of the  
RT interleukin-4 receptor.";  
RL Science 262:1880-1883(1993).  
RN [6]  
RP IDENTIFICATION AS A IL-7R SUBUNIT.  
RX MEDLINE; 94090316.  
RA Noguchi M., Nakamura Y., Russell S.M., Ziegler S.F., Tsang M., Cao  
RA Leonard W.J.;  
RT "Interleukin-2 receptor gamma chain: a functional component of the  
RT interleukin-7 receptor.";  
RL Science 262:1877-1880(1993).  
RN [7]  
RP 3D-STRUCTURE MODELING OF 57-248.  
RX MEDLINE; 95111955.  
RA Bamorough P., Hedgecock C.J., Richards W.G.;  
RT "The interleukin-2 and interleukin-4 receptors studied by molecular  
RT modelling.";  
RL Structure 2:839-851(1994).  
RN [8]  
RP VARIANTS XSCID PHE-115; CYS-240 AND ILE-241.  
RX MEDLINE; 94130970.  
RA Disanto J.P., Dautry-Varsat A., Certain S., Fischer A.,  
RA de Saint Basile G.;  
RT "Interleukin-2 (IL-2) receptor gamma chain mutations in X-linked  
RT severe combined immunodeficiency disease result in the loss of  
RT high-affinity IL-2 receptor binding.";  
RL Eur. J. Immunol. 24:475-479(1994).  
RN [9]  
RP VARIANT XSCID LYS-68.  
RX MEDLINE; 94375038.  
RA Markiewicz S., Subtil A., Dautry-Varsat A., Fischer A.,  
RA de Saint Basile G.;  
RT "Detection of three nonsense mutations and one missense mutation in  
RT the interleukin-2 receptor gamma chain gene in SCIDX1 that  
RT differently affect the mRNA processing.";  
RL Genomics 21:291-293(1994).  
RN [10]  
RP VARIANT XSCID HIS-162.  
RX MEDLINE; 94350093.  
RA Ishii N., Asao H., Kimura Y., Takeshita T., Nakamura M., Tsuchiya S.,  
RA Konno T., Maeda M., Uchiyama T., Sugamura K.;  
RT "Impairment of ligand binding and growth signaling of mutant IL-2  
RT receptor gamma-chains in patients with X-linked severe combined  
RT immunodeficiency.";  
RL J. Immunol. 153:1310-1317(1994).  
RN [11]  
RP VARIANT XSCID ASN-39.  
RX MEDLINE; 95023932.  
RA Disanto J.P., Rieux-Laucat F., Dautry-Varsat A., Fischer A.,  
RA de Saint Basile G.;  
RT "Defective human interleukin 2 receptor gamma chain in an atypical X  
RT chromosome-linked severe combined immunodeficiency with peripheral  
RT cells.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:9466-9470(1994).  
RN [12]  
RP VARIANTS XSCID CYS-226 AND HIS-226.  
RX MEDLINE; 95097841.  
RA Pepper A.E., Buckley R.H., Small T.N., Puck J.M.;  
RT "Two mutational hotspots in the interleukin-2 receptor gamma chain  
RT gene causing human X-linked severe combined immunodeficiency.";  
RN

RESULT	2	
ID	CYRG_HUMAN	STANDARD; PRT; 369 AA.
AC	P31785;	
DT	01-JUL-1993	(Rel. 26, Created)
DT	01-JUL-1993	(Rel. 26, Last sequence update)
DT	15-DEC-1998	(Rel. 37, Last annotation update)
DE	CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C) (INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64) (CD132 ANTIGEN).	
GN	IL2RG.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
OC	Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
RP	[1]	
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.	
RX	MEDLINE; 92335883.	
RA	Takeshita T., Asao H., Ohtani K., Ishii N., Kumaki S., Tanaka N.,	
RA	Munakata H., Nakamura M., Sugamura K.;	
RT	"Cloning of the gamma chain of the human IL-2 receptor.";	
RL	Science 257:379-382(1992).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=LIVER;	
RX	MEDLINE; 93293887.	
RA	Nouchi M., Adelstein S., Cao X., Leonard W.J.;	
RT	"Characterization of the human interleukin-2 receptor gamma chain	
RT	gene.";	
RL	J. Biol. Chem. 268:13601-13608(1993).	
RN	[3]	
RP	SEQUENCE FROM N.A., AND VARIANTS ASP-114 AND ASN-153.	
RX	MEDLINE; 94004847.	
RA	Puck J.M., Deschenes S.M., Porter J.C., Dutra A.S., Brown C.J.,	
RA	Willard H., Henchorn P.S.;	
RT	"The interleukin-2 receptor gamma chain maps to Xq13.1 and is mutated	
RT	in X-linked severe combined immunodeficiency, SCIDX1.";	
RL	Hum. Mol. Genet. 2:1099-1104(1993).	
RN	[4]	

Am. J. Hum. Genet. 57:564-571(1995).  
 [13] VARIANT XSCID SER-183.  
 RA MEDLINE; 96013903.  
 RA Clark P.A., Lester T., Genet S., Jones A.M., Hendriks R.,  
 RA Levensky R.L., Kinnon C.; causing X-linked severe combined  
 RT immunodeficiency in the IL-2R gamma chain gene by single-strand  
 RT conformation polymorphism analysis.  
 RL Hum. Genet. 96:427-432(1995).  
 [14]  
 RP VARIANT XSCID GLN-HIS-TRP INS-237.  
 RA MEDLINE; 95164726.  
 RA Puck J.M., Pepper A.E., Bedard P.-M., Laframboise R.;  
 RA "Female germ line mosaicism as the origin of a unique IL-2 receptor  
 RT gamma-chain mutation causing X-linked severe combined  
 RT immunodeficiency".  
 RL J. Clin. Invest. 95:895-899(1995).  
 [15]  
 RP VARIANT XSCID GLN-271.  
 RA MEDLINE; 95190013.  
 RA Schmalstieg F.C., Leonard W.J., Noguchi M., Berg M., Rudloff H.E.,  
 RA Denney R.M., Dave S.K., Brooks E.G., Goldman A.S.;  
 RA "Missense mutation in exon 7 of the common gamma chain gene causes a  
 RT moderate form of X-linked combined immunodeficiency".  
 RL J. Clin. Invest. 95:1169-1173(1995).  
 [16]  
 RP VARIANT XSCID ARG-115.  
 RA MEDLINE; 97042245.  
 RA Stephan V., Wahn V., le Deist F., Dirksen U., Broker B.,  
 RA Muller-Eckstein I., Horneff G., Schrotten H., Fischer A.,  
 RA de Saint Basile G.;  
 RA "Atypical X-linked severe combined immunodeficiency due to possible  
 RT spontaneous reversion of the genetic defect in T cells".  
 RL New Engl. J. Med. 335:1563-1567(1996).  
 [17]  
 RP VARIANT XSCID GLN-285.  
 RA MEDLINE; 97295088.  
 RA Jones A.M., Clark P.A., Katz F., Genet S., McMahon C., Alterman L.,  
 RA Cant A., Kinnon C.;  
 RA "B-cell-negative severe combined immunodeficiency associated with a  
 RT common gamma chain mutation".  
 RL Hum. Genet. 99:677-680(1997).  
 [18]  
 RP VARIANT XSCID CYS-222.  
 RA MEDLINE; 98064061.  
 RA Sharie N., Shahar M., Roifman C.M.;  
 RA "An interleukin-2 receptor gamma chain mutation with normal thymus  
 RT morphology".  
 RL J. Clin. Invest. 100:3036-3043(1997).  
 CC -!- FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF  
 CC INTERLEUKINS.  
 CC -!- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND  
 CC PROBABLY ALSO THE IL-13 RECEPTORS.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- DISEASE: DEFECTS IN IL2RG ARE THE CAUSE OF A SEVERE COMBINED  
 CC OR X-LINKED SEVERE COMBINED IMMUNODEFICIENCY DISEASE (XSCID).  
 CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD132 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd132.htm".  
 CC -!- DATABASE: NAME=IL2RGbase; NOTE=X-linked SCID mutation database;  
 CC WWW="http://www.nhgri.nih.gov/DIR/IGT/SCID/IL2RGbase.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; D11086; BAA01857.1; -.  
 DR EMBL; L12183; AAA59145.1; -.  
 DR EMBL; L12178; AAA59145.1; JOINED.  
 DR EMBL; L12176; AAA59145.1; JOINED.  
 DR EMBL; L12177; AAA59145.1; JOINED.  
 DR EMBL; L12179; AAA59145.1; JOINED.  
 DR EMBL; L12180; AAA59145.1; JOINED.  
 DR EMBL; L12181; AAA59145.1; JOINED.  
 DR EMBL; L12182; AAA59145.1; JOINED.  
 DR EMBL; L19546; AAC37524.1; -.  
 DR PIR; A42565; A42565.  
 DR PDB; 1ILM; 26-JAN-95.  
 DR PDB; 1ILN; 26-JAN-95.  
 DR MIM; 300400; -.  
 DR MIM; 308380; -.  
 DR PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
 DR PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.  
 DR PFAM; PF00041; fn3; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Signal; Disease mutation;  
 ...  
 Note: remainder of annotations omitted.  
 Query Match 11.0%; Score 195; DB 1; Length 369;  
 Best Local Similarity 25.9%; Pred. No. 3.99e-20;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 59 EVOCFVNVVEYMCNTWNSSEPOPTNLTHYWKNSDNDKVKCSHYLFSEETISGCOLQ 118  
 QY 31 QIQIIIFNLETVQVTVNASKYSR-TNLTFRYR-NGD-EAYDOCTNYLLQEGHTSGCLLD 87  
 Db 119 KKEIHLYQTFVVOLOQPPRRQATQMLQNLVWPAPENLTLHLKLSQLELNWNRF 178  
 QY 88 AEQRDDILFESIR-NGTH-EVFTASRWV-VYIL-KPSSPKHVRP-SWHQDAVTVTCSDL 142  
 Db 179 LNHCLHLVQYRTDWDHSWTEQSVYRHKFSLPSVDQKRYTFVRSR-FNPLGSAQHW 237  
 QY 143 YGD-LLVEVQYRSPFDEW-QSKQENTCNVTIEGLDAEKCYSEWVRVKAMEDVYGPDTYP 200  
 Db 238 SEWSHPZHW 246  
 QY 201 SDWSEVTCW 209  
 ...  
 RESULT 3  
 ID CYRG\_MOUSE STANDARD; PRT; 369 AA.  
 AC P34902;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C)  
 DE (INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64).  
 GN IL2RG.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93277575.  
 RA Kumaki S., Kondo M., Takeshita T., Asao H., Nakamura M., Sugamura K.;  
 RT "Cloning of the mouse interleukin 2 receptor gamma chain:  
 RT demonstration of functional differences between the mouse and human  
 RT receptors".  
 RL Biochem. Biophys. Res. Commun. 193:356-363(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SRAIN-CBA/CA.  
 RA MEDLINE; 93391374.  
 RA Cao X., Kozak C.A., Liu Y.J., Noguchi M., O'Connell E., Leonard W.J.;  
 RT "Characterization of cDNAs encoding the murine interleukin 2 receptor  
 RT (IL-2R) gamma chain: chromosomal mapping and tissue specificity of  
 RT IL-2R gamma chain expression".  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:8464-8468(1993).  
 RN [3]

RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93366191.  
 RA Kobayashi N., Nakagawa S., Minami Y., Taniguchi T., Kono T.;  
 RT "Cloning and sequencing of the cDNA encoding a mouse IL-2 receptor  
 RL gamma.;"  
 RN Gene 130:303-304(1993).  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 95104285.  
 RA Disanto J.P., Certain S., Wilson A., Macdonald H.R., Avner P.,  
 RA Fischer A., de Saint Basile G.;  
 RT "The murine interleukin-2 receptor gamma chain gene: organization,  
 RL chromosomal localization and expression in the adult thymus.;"  
 RN Eur. J. Immunol. 24:3014-3018(1994).  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=B6.S;  
 RX MEDLINE; 96341745.  
 RA Chiu R.K., Droll A., Cooper D.L., Dougherty S.T., Dirks J.F.,  
 RA Dougherty G.J.;  
 RT "Molecular mechanisms regulating the hyaluronan binding activity of  
 RT the adhesion protein CD44.;"  
 RL J. Neurooncol. 26:231-239(1995).  
 CC -1- FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF  
 CC INTERLEUKINS.  
 CC -1- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND  
 CC PROBABLY ALSO THE IL-13 RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D13821; BAA02974.1; -;  
 DR EMBL; U21795; BAA04279.1; -;  
 DR EMBL; D13565; BAA02760.1; -;  
 DR EMBL; L20048; AAB39286.1; -;  
 DR EMBL; S75852; AAB32904.1; -;  
 DR EMBL; S75844; AAB32904.1; JOINED.  
 DR EMBL; S75845; AAB32904.1; JOINED.  
 DR EMBL; S75847; AAB32904.1; JOINED.  
 DR EMBL; S75848; AAB32904.1; JOINED.  
 DR EMBL; S75849; AAB32904.1; JOINED.  
 DR EMBL; S75850; AAB32904.1; JOINED.  
 DR EMBL; S75851; AAB32904.1; JOINED.  
 DR EMBL; X75337; CAA53085.1; -;  
 DR PIR; JN0592; JN0592; -;  
 DR PIR; JN0775; JN0775; -;  
 DR HSP; P31785; IL2RG.  
 DR MGD; MGI:96551; IL2RG.  
 DR PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
 DR PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.  
 DR PFAM; PF00041; fn3; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 369  
 FT DOMAIN 23 263  
 FT TRANSMEM 264 284  
 FT DOMAIN 285 369  
 FT DOMAIN 151 250  
 FT CYTOPLASMIC (POTENTIAL).  
 FT FIBRONECTIN TYPE-III.  
 FT POTENTIAL.  
 FT DISULFID 62 72  
 FT DISULFID 102 115  
 FT CARBOHYD 71 71  
 FT POTENTIAL.  
 FT CARBOHYD 75 75  
 FT POTENTIAL.  
 FT CARBOHYD 84 84  
 FT POTENTIAL.  
 FT CARBOHYD 96 96  
 FT POTENTIAL.  
 FT CARBOHYD 159 159  
 FT POTENTIAL.

FT CARBOHYD 164 164 POTENTIAL.  
 SQ SEQUENCE 369 AA; 42241 MW; CB2D5AB459077AC7 CRC64;  
 Query Match 10.6%; Score 188; DB 1; Length 369;  
 Best Local Similarity 26.8%; Pred. No. 1.16e-18;  
 Matches 51; Conservative 52; Mismatches 75; Indels 12; Gaps 11;  
 Db 59 EVQCFVFNEMNCTWNSSEPOATNLTHRYKVKVSNNTFQECSHYLFSEKTSQCIQ 118  
 QY 31 QQIIIFNLETQVTVNASKYSR-TNLTFFHYRFN-GDE-AYDOCTNYLLQGHSTGCLLD 87  
 Db 119 KEDIQIQYQFVVOLOQPOKPORRAVOKLNQNLVPRAPENLTLSNSESOLELRWKSRL 178  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYLL-KFSSPRHVRFS-WHQDAVTW--TCSD 140  
 Db 179 IKERCLQYLVQVRSNRDRSWTELIVNHEPRFSLPSDELKRYTFRVRSRY-NPICGSSQ 237  
 QY 141 LSYGDLLEYQVRSPPDTEW-QSKQENTCNVTIEGLDAEKCYSEFWVRVKAMEDYVGPDTY 199  
 Db 238 WSKWSQPVHW 247  
 QY 200 PSDWSEVTCW 209  
 RESULT 4  
 ID CYRG.BOVIN STANDARD; PRT; 379 AA.  
 AC Q95118;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C)  
 DE (INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64).  
 GN IL2RG.  
 OS Bos taurus (Bovine).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 CC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovinae; Bos.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96268473.  
 RA Yoo J., Stone R.T., Solinas-Toldo S., Fries R., Beattie C.W.;  
 RT "Cloning and chromosomal mapping of bovine interleukin-2 receptor  
 RL gamma gene.;"  
 RL DNA Cell Biol. 15:453-459(1996).  
 CC -1- FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF  
 CC INTERLEUKINS.  
 CC -1- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND  
 CC PROBABLY ALSO THE IL-13 RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U33748; BAB07812.1; -;  
 DR HSP; P31785; IL2N.  
 DR PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
 DR PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.  
 DR PFAM; PF00041; fn3; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 379  
 FT CYTOKINE RECEPTOR COMMON GAMMA CHAIN.  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 270 290  
 FT DOMAIN 291 379  
 FT CYTOPLASMIC (POTENTIAL).  
 FT FIBRONECTIN TYPE-III.  
 FT POTENTIAL.  
 FT DISULFID 68 78

FT CHAIN 23 878 INTERLEUKIN-3 RECEPTOR CLASS II BETA  
 FT DOMAIN 23 440 CHAIN.  
 FT TRANSMEM 441 462 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 463 878 POTENTIAL.  
 FT DISULFID 39 49 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 78 95 BY SIMILARITY.  
 FT DISULFID 254 264 BY SIMILARITY.  
 FT CARBOHYD 293 310 BY SIMILARITY.  
 FT CARBOHYD 62 62 POTENTIAL.  
 FT CARBOHYD 350 350 POTENTIAL.  
 SQ SEQUENCE 878 AA; 97195 MW; 8EBC9092ADC24D56 CRC64;

Query Match 10.5%; Score 187; DB 1; Length 379;  
 Best Local Similarity 29.3%; Pred. No. 1.87e-18; Indels 19; Gaps 17;  
 Matches 61; Conservative 49; Mismatches 79; Indels 19; Gaps 17;  
 Db 66 VQCFVENVEYMTNCTWSSSSPQPNLTHYGYRNFNGDKLQECGHLFSEGTSCWF- 124  
 Qy 32 IQIYFNLEIVQVWNAKYSRTN-LTFHY--R-FNGDEAYDOCTNYLQEGHTSCGLD 87  
 Db 125 G-KKEIRLYETFFVQLODPREHRKQPKMLQDLOLVIPWAPENLTNLNLSFQLELSWSN 183  
 Qy 88 AEQDDILY--FSIR-NGTHPVFTASRWV-YY-L-KPSPKHVRSWHQD-AVTVCSD 140  
 Db 184 -RYLDHCHLEHLVQYRDRSWSQSDVDRHSFSLPSVDAQKLYTFVRSRY-NPLCGSA 241  
 Qy 141 LSYGDL-L-YEQYRSPFTDTEW-QSKQENTCNVTIEGLDAEKCYSFVVRVKAEDVYGP 197  
 Db 242 QHWSDSYPIHWGNTSKEN-IENPENP 268  
 Qy 198 TYPDSWSEVTCWGEIRDACAETPTP 225

RESULT 5  
 ID IL3B\_MOUSE STANDARD; PRT; 878 AA.  
 AC P26954;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-FEB-2000 (Rel. 39, Last annotation update)  
 DE INTERLEUKIN-3 RECEPTOR CLASS II BETA CHAIN PRECURSOR (COLONY  
 DE STIMULATING FACTOR 2 RECEPTOR, BETA 2 CHAIN).  
 GN CSF2RB2 OR A12CA OR IL3RB2 OR IL3R.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90117145.  
 RA Itoh N., Yonehara S., Schreurs J., Gorman D.M., Maruyama K., Ishii A.,  
 RA Yahara I., Arai K., Miyajima A.;  
 RT "Cloning of an interleukin-3 receptor gene: a member of a distinct  
 RT receptor gene family."  
 RL Science 247:324-327(1990).  
 CC -1- FUNCTION: IN MOUSE THERE ARE TWO CLASSES OF HIGH-AFFINITY IL-3  
 CC RECEPTORS. ONE CONTAINS THIS IL-3-SPECIFIC BETA CHAIN AND THE  
 CC OTHER CONTAINS THE BETA CHAIN ALSO SHARED BY HIGH-AFFINITY IL-5  
 CC AND GM-CSF RECEPTORS.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M29855; AAA39295.1;  
 CC DR PIR; A40091; A40091.  
 CC DR MGD; MGI:1339760; CSF2RB2.  
 CC DR PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
 CC DR PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; 1.  
 CC DR PFAM; PF00041; fn3; 2.  
 CC KW Receptor; Transmembrane; Glycoprotein; Signal.  
 CC SIGNAL 1 22  
 CC POTENTIAL.

RESULT 6  
 ID CYRB\_HUMAN STANDARD; PRT; 897 AA.  
 AC P32927;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE CYTOKINE RECEPTOR COMMON BETA CHAIN PRECURSOR (CD131 ANTIGEN).  
 GN CSF2RB OR IL5RB OR IL3RB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91088571.  
 RA Hayashida K., Kitamura T., Gorman D.M., Arai K., Yokota T.,  
 RA Miyajima A.;  
 RT "Molecular cloning of a second subunit of the receptor for human  
 RT granulocyte/macrophage colony-stimulating factor (GM-CSF):  
 RT reconstitution of a high-affinity GM-CSF receptor."  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9655-9659(1990).  
 RN [2]  
 RP REVISION TO 454.  
 RA Kitamura T.;  
 RA Submitted (FEB-1991) to the EMBL/GenBank/DBJ databases.  
 RL -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR INTERLEUKIN-3, INTERLEUKIN-5  
 CC AND GRANULOCYTE-MACROPHAGE COLONY-STIMULATING FACTOR.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. THE BETA  
 CC CHAIN IS COMMON TO THE IL-3, IL-5 AND GM-CSF RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide cdw131 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cdw131.htm".  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M59941; AAA18171.1;  
 CC DR PIR; A39255; A39255.  
 CC DR MIM; 138981;  
 CC DR PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 2.  
 CC DR PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; 1.





CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X95302; CAA64617.1; -

CC EMBL; U70981; AAB17170.1; -

CC EMBL; Y08768; CAA70021.1; -

CC MIM; 300130; -

CC PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; 1.

CC PFAM; PF00041; fn3; 1.

CC Receptor; Transmembrane; Glycoprotein; Signal.

CC SIGNAL 1 26

CC CHAIN 27 380

CC INTERLEUKIN-13 RECEPTOR ALPHA-2 CHAIN.

CC DOMAIN 27 343

CC EXTRACELLULAR (POTENTIAL).

CC TRANSMEM 344 363

CC POTENTIAL.

CC DOMAIN 364 380

CC CYTOPLASMIC (POTENTIAL).

CC DISULFID 145 155

CC BY SIMILARITY.

CC DISULFID 194 197

CC BY SIMILARITY.

CC CARBOHYD 115 115

CC POTENTIAL.

CC CARBOHYD 215 215

CC POTENTIAL.

CC CARBOHYD 250 290

CC POTENTIAL.

CC CARBOHYD 239 299

CC POTENTIAL.

CC QUERY SEQUENCE 380 AA; 44176 MW; 3C6ACB1B5562C887 CRC64;

Query Match 5.8%; Score 102; DB 1; Length 380;

Best Local Similarity 28.2%; Pred. No. 1.63e-02;

Matches 20; Conservative 21; Mismatches 24; Indels 6; Gaps 5;

Db 129 WIS-PQIGPETHVQMDVYNNQYLCSWKPGIGVLLDTNLYNFYWEGLDHAL-QCYD 186

QY 18 WMAQGGGAEGVQ-IGIIFNLETQVVTWNSK--YSRTNLTFFHFNCG-DEAYDCCIN 73

Db 187 YIKADGNIGCC 197

QY 74 YLLQEGHTSGC 84

RESULT 9

ID GHR HUMAN STANDARD; PRT; 638 AA.

AC P10912;

DT 01-JUL-1989 (Rel. 11, Created)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE GROWTH HORMONE RECEPTOR PRECURSOR (GHR RECEPTOR) (SERUM BINDING PROTEIN).

GN GHR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

[1]

RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

CC TISSUE=LIVER;

CC MEDLINE; 88065896.

CC Leung D.W., Spencer S.A., Cachianes G., Hammonds R.G., Collins C., Henzel W.J., Barnard R., Waters M.J., Wood W.I.;

CC "Growth hormone receptor and serum binding protein: purification, cloning and expression."

CC Nature 330:537-543(1987).

[2]

CC SEQUENCE FROM N.A.

CC MEDLINE; 90046742.

CC Godowski P.J., Leung D.W., Meacham L.R., Galgani J.P., Hellmiss R., Keret R., Rotwein P.S., Parks J.S., Larson Z., Wood W.I.;

CC "Characterization of the human growth hormone receptor gene and demonstration of a partial gene deletion in two patients with Laron-type dwarfism."

CC Proc. Natl. Acad. Sci. U.S.A. 86:8083-8087(1989).

[3]

CC DISULFIDE BONDS

CC MEDLINE; 90153957.

CC Fuh G., Mulkerin M.G., Bass S., McFarland N., Brochier M., Bourrel J.H., Light D.R., Wells J.A.;

CC "The human growth hormone receptor. Secretion from Escherichia coli and disulfide bonding pattern of the extracellular binding domain."

CC RT

J. Biol. Chem. 265:3111-3115(1990).

[4]

CC MEDLINE; 89384829.

CC Amselem S., Duquesnoy P., Attree O., Novelli G., Bousnina S., Postelvinay M.-C., Goossens M.;

CC "Laron dwarfism and mutations of the growth hormone-receptor gene."

CC New Engl. J. Med. 321:989-995(1989).

[5]

CC VARIANTS LARON DWARFISM.

CC MEDLINE; 93278381.

CC Amselem S., Duquesnoy P., Duriez B., Dastot F., Sorbier M.-L., Valleix S., Goossens M.;

CC "Spectrum of growth hormone receptor mutations and associated haplotypes in Laron syndrome."

CC Hum. Mol. Genet. 2:355-359(1993).

[6]

CC VARIANTS LARON DWARFISM HIS-170.

CC MEDLINE; 94185645.

CC Dusquesnoy P., Sobrier M.-L., Duriez B., Dastot F., Buchanan C.R., Savage M.O., Preece M.A., Craescu C.T., Blouquit Y., Goossens M., Anselem S.;

CC "A single amino acid substitution in the exoplasmic domain of the human growth hormone (GH) receptor confers familial GH resistance (Laron syndrome) with positive GH-binding activity by abolishing receptor homodimerization."

CC EMBO J. 13:1386-1395(1994).

[7]

CC VARIANTS IDIOPATHIC SHORT STATURE LYS-62; CYS-179 AND ASP-242.

CC MEDLINE; 96013502.

CC Godard A.D., Cuvellier R., Luch S.-M., Clackson T., Attie K.M., Gessendheit N., Rundle A.C., Wells J.A., Carlsson L.M.S.;

CC "Mutations of the growth hormone receptor in children with idiopathic short stature. The Growth Hormone Insensitivity Study Group."

CC New Engl. J. Med. 333:1093-1098(1995).

[8]

CC X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 19-256.

CC MEDLINE; 92196577.

CC de Vos A.M., Gitsch M., Kossiakoff A.A.;

CC "Human growth hormone and extracellular domain of its receptor: crystal structure of the complex."

CC Science 255:306-312(1992).

[9]

CC X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 19-256.

CC MEDLINE; 97113023.

CC Sundstroem M., Lundqvist T., Rodin J., Giebel L.B., Milligan D., Norstedt G.;

CC "Crystal structure of an antagonist mutant of human growth hormone, G120R, in complex with its receptor at 2.9-A resolution."

CC J. Biol. Chem. 271:32197-32203(1996).

CC -1- FUNCTION: THIS IS A RECEPTOR FOR PITUITARY GLAND GROWTH HORMONE.

CC -1- SUBUNIT: HOMODIMER.

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -1- DISEASE: DEFICIENCY IN GHR IS THE CAUSE OF PITUITARY DWARFISM II (LARON-TYPE PITUITARY DWARFISM OR LARON SYNDROME (LS)). IT ALSO CAUSES IDIOPATHIC SHORT STATURE.

CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.

CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed (usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))).

CC -----

CC EMBL; X06562; CAA29808.1; -

CC EMBL; M28466; AAA52555.1; -

CC EMBL; M28458; AAA52555.1; JOINED.

CC EMBL; M28459; AAA52555.1; JOINED.

CC EMBL; M28460; AAA52555.1; JOINED.

CC EMBL; M28461; AAA52555.1; JOINED.

CC

DR EMBL; M28462; AAA52555.1; JOINED.  
 DR EMBL; M28463; AAA52555.1; JOINED.  
 DR EMBL; M28464; AAA52555.1; JOINED.  
 DR EMBL; M28465; AAA52555.1; JOINED.  
 DR PIR; S04530; S04530.  
 DR PIR; S03991; A33991.  
 DR PDB; 3HR; 30-APR-94.  
 DR PDB; 1HW; 19-NOV-97.  
 DR PDB; 1WH; 19-NOV-97.  
 DR PDB; 1AXI; 28-JAN-98.  
 DR PDB; 1A22; 29-APR-98.  
 DR MIM; 600946; .  
 DR MIM; 262500; .  
 DR PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
 DR PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.  
 DR PFAM; PF00041; fn3; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Signal; 3D-structure;  
 Dwarfishm; Disease mutation.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 638 GROWTH HORMONE RECEPTOR.  
 FT DOMAIN 19 264 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 265 288 POTENTIAL.  
 FT DOMAIN 289 638 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 145 252 FIBRONECTIN TYPE-III.  
 FT DISULFID 56 66  
 FT DISULFID 101 112  
 FT DISULFID 126 140  
 FT DISULFID 126 140  
 FT CARBOHYD 46 46  
 FT CARBOHYD 115 115  
 FT CARBOHYD 156 156  
 FT CARBOHYD 161 161  
 FT CARBOHYD 200 200  
 FT VARIANT 62 62  
 FT VARIANT 89 89  
 FT VARIANT 114 114  
 FT VARIANT 143 143  
 FT VARIANT 162 162  
 FT VARIANT 170 170  
 FT VARIANT 179 179  
 FT VARIANT 229 229  
 FT VARIANT 242 242  
 FT CONFLICT 544 544  
 FT STRAND 53 58  
 FT STRAND 64 68  
 FT STRAND 82 88  
 FT STRAND 99 100  
 FT TURN 104 107  
 FT TURN 109 110  
 FT STRAND 111 114  
 FT TURN 116 117  
 FT STRAND 124 131  
 FT TURN 132 133  
 FT STRAND 134 142  
 FT HELIX 143 145  
 FT STRAND 147 147  
 FT STRAND 153 162  
 FT TURN 164 165  
 FT STRAND 168 176  
 FT TURN 179 180  
 FT TURN 183 186  
 FT STRAND 190 198  
 FT TURN 199 200

FT STRAND 205 206  
 FT STRAND 210 210  
 FT STRAND 214 221  
 FT TURN 222 223  
 FT STRAND 225 234  
 FT STRAND 247 250  
 SQ SEQUENCE 638 AA; 71499 MW; EAF77EAD84787822 CRC64;  
 Query Match 5.6%; Score 100; DB 1; Length 638;  
 Best Local Similarity 32.7%; Pred. No. 3.35e-02;  
 Matches 17; Conservative 11; Mismatches 22; Indels 2; Gaps 2;  
 Db 190 LEYELQYKEVNETKWKMDPILTTSTVPYSLKVDKEYEVVRVSKORNSGNYG 241  
 QY 146 LLIEVOYRSPFDTEWQSKOEN-TCNVTIEGLDAEKCYSFWRVKAMEDV-YG 195  
 RESULT 10  
 ID YEO6\_YEAST STANDARD; PRT; 302 AA.  
 AC P40049;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HYPOTHETICAL 33.5 KD PROTEIN IN PTP3-ILV1 INTERGENIC REGION  
 DE PRECURSOR.  
 GN YER076C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB9712;  
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunkle-Smith S.,  
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,  
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,  
 RA Petel F.X., Roberts D., Seh P., Schramm S., Shogren T., Smith V.,  
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: TO YEAST KILLER TOXIN KHR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U18839; AAB64631.1;  
 KW Hypothetical protein; Glycoprotein; Signal.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 302 HYPOTHETICAL PROTEIN YER076C.  
 FT CARBOHYD 65 65  
 FT CARBOHYD 86 86  
 FT CARBOHYD 93 93  
 FT CARBOHYD 220 220  
 FT CARBOHYD 231 231  
 FT CARBOHYD 231 231  
 SQ SEQUENCE 302 AA; 33466 MW; 43E3AC5FA2BF378F CRC64;  
 Query Match 5.5%; Score 97; DB 1; Length 302;  
 Best Local Similarity 22.8%; Pred. No. 9.72e-02;  
 Matches 18; Conservative 21; Mismatches 33; Indels 7; Gaps 7;  
 Db 194 CGSQEFTNIFDQEGWSLFLVKTW-ST-NSSCDITASE-GNLTC-AVRVSVSSMIN-HCK 248  
 QY 138 CSDLGYGDLLEYVQY-RSPFDEWQSKOENCNVTIEGLDAEKCYSFWRVKAMEDVYGP 196  
 Db 249 TAFCVTYSHGDSW-RAELR 266  
 QY 197 DTYPDSWSEVTCWQGEIR 215

US-09-376-430-2-03.rsp

Thu May 11 06:49:30 2000

SEQUENCE FROM N.A.  
STRAIN-LANDRACE-YORKSHIRE; TISSUE=LIVER;  
RC MEDLINE; 91057155.  
RA Cloff J.A., Wang X., Kopchick J.J.;  
RT "Porcine growth hormone receptor cDNA sequence."  
RL Nucleic Acids Res. 18:6451-6451(1990).  
CC -!- FUNCTION: THIS IS A RECEPTOR FOR PITUITARY GLAND GROWTH HORMONE.  
CC -!- SUBUNIT: HOMODIMER.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X54429; CAA38301.1; -  
CC PIR; S12136; S12136.  
CC HSP; P10912; 1A22.  
CC PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.  
CC PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
CC PFAM; PF00041; fn3; 1.  
CC  
CC Receptor; Transmembrane; Glycoprotein; Signal.  
CC SIGNAL  
CC CHAIN 1  
CC FT DOMAIN 19 638  
CC FT DOMAIN 19 264  
CC FT TRANSMEM 265 288  
CC FT DOMAIN 289 638  
CC FT DOMAIN 145 252  
CC FT DISULFID 56 66  
CC FT DISULFID 101 112  
CC FT DISULFID 126 140  
CC FT CARBOHYD 146 46  
CC FT CARBOHYD 115 115  
CC FT CARBOHYD 156 156  
CC FT CARBOHYD 161 161  
CC FT CARBOHYD 200 200  
CC SQ SEQUENCE 638 AA; 71145 MW; BC7C66536F4DFF97 CRC64;  
Query Match 5.4%; Score 96; DB 1; Length 638;  
Best Local Similarity 28.8%; Pred. No. 1.38e-01;  
Matches 15; Conservative 15; Mismatches 20; Indels 2; Gaps 2;  
Db 190 LYELOKKEVNETQWKMDPVLSTVPVSLRLDKVEYEVRSQRNSEKYG 241  
QY 146 LLYEVOIRSPFDIEMOSKQEN-TCNVTIEGLDAEKCYSFWRVKAME-DVYG 195

SEQUENCE FROM N.A.  
STRAIN-AH65;  
RC MEDLINE; 96417863.  
RA Pepe C.M., Eklund M.W., Strom M.S.;  
RT "Cloning of an Aeromonas hydrophila type IV pilus biogenesis gene  
RT cluster: complementation of pilus assembly functions and  
RT characterization of a type IV leader peptidase/N-methyltransferase  
RT required for extracellular protein secretion."  
RL Mol. Microbiol. 19:857-869(1996).  
CC -!- FUNCTION: CLEAVES TYPE-4 FIMBRIAL LEADER SEQUENCE AND METHYLATES  
CC THE N-TERMINAL (GENERALLY PHE) RESIDUE. PROCESSES THE TAPA PILLIN  
CC PRECURSOR DURING MEMBRANE TRANSLOCATION.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE  
CC (PROBABLE).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C20; ALSO KNOWN AS TYPE  
CC III LEADER PEPTIDASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U20255; AAC43998.1; -  
CC PFAM; PF01478; Peptidase\_C20; 1.  
CC PRINTS; PR00864; PREPILNPTASE.  
CC Hydrolase; Transport; Transmembrane; Inner membrane.  
CC TRANSMEM 14 34  
CC TRANSMEM 106 126  
CC TRANSMEM 130 150  
CC TRANSMEM 161 181  
CC TRANSMEM 185 205  
CC TRANSMEM 232 252  
CC TRANSMEM 261 281  
CC SQ SEQUENCE 290 AA; 32308 MW; EA584F277041A99B CRC64;  
Query Match 5.4%; Score 96; DB 1; Length 290;  
Best Local Similarity 54.5%; Pred. No. 1.38e-01;  
Matches 12; Conservative 6; Mismatches 3; Indels 1; Gaps 1;  
Db 163 LPLWGLLNLGLGFFVSLGDA 184  
QY 4 LVLLWGAVF-LLGWVALGOG 24

RESULT 12  
ID GHR.PIG STANDARD; PRT; 638 AA.  
AC P19756;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE GROWTH HORMONE RECEPTOR PRECURSOR (GH RECEPTOR) (SERUM BINDING  
DE PROTEIN).  
GN GHR.  
OS Sus scrofa (Pig).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
RN [1]

CC -!- SUBUNIT: HOMODIMER.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF015252; AAB67613.1; -  
 CC PIR: S08544; S08544.  
 CC HSSP: P10912; IA22.  
 CC PROSITE: PS00340; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.  
 CC DR PFAM: PF00041; fn3; 1.  
 CC KW Receptor; Transmembrane; Glycoprotein; Signal.  
 CC FT SIGNAL 1 18  
 CC FT CHAIN 19 638 GROWTH HORMONE RECEPTOR.  
 CC FT DOMAIN 19 264 EXTRACELLULAR (POTENTIAL).  
 CC FT TRANSMEM 265 288 POTENTIAL.  
 CC FT DOMAIN 289 638 CYTOPLASMIC (POTENTIAL).  
 CC FT DOMAIN 145 252 FIBRONECTIN TYPE-III.  
 CC FT DISULFID 56 66 BY SIMILARITY.  
 CC FT DISULFID 101 112 BY SIMILARITY.  
 CC FT DISULFID 126 140 BY SIMILARITY.  
 CC FT CARBOHYD 46 46 POTENTIAL.  
 CC FT CARBOHYD 115 115 POTENTIAL.  
 CC FT CARBOHYD 156 156 POTENTIAL.  
 CC FT CARBOHYD 161 161 POTENTIAL.  
 CC FT CARBOHYD 200 200 POTENTIAL.  
 CC SQ SEQUENCE 638 AA; 71076 MW; E0SCCEID794624C CRC64;

Query Match 5.3%; Score 94; DB 1; Length 638;  
 Best Local Similarity 28.8%; Pred. No. 2.75e-01;  
 Matches 15; Conservative 14; Mismatches 21; Indels 2; Gaps 2;

Db 190 LEVELQYKEVNETOKMMPVLSTSPVSLRLDKYEVVRQRSSSEKYG 241  
 QY 146 LLYEYQVSPFTDQWQSKOEN-TCNVTIEGLDAEKCYFWVRKAME-DVYG 195

RESULT 14  
 ID GLNA\_PANAR STANDARD; PRT; 361 AA.  
 AC Q04831;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE GLUTAMINE SYNTHETASE (EC 6.3.1.2) (GLUTAMATE--AMMONIA LIGASE).  
 OS Panulirus argus (Spiny lobster).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Palinura;  
 OC Palinuridae; Panulirus.  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC TISSUE-OLFACTORY ORGAN;  
 CC MEDLINE; 93314973.  
 CC RA Trapido-Rosenthal H.G., Linser P.J., Greenberg R.M., Gleeson R.A.,  
 CC Carr W.E.,  
 CC "cDNA clones from the olfactory organ of the spiny lobster encode a  
 CC protein related to eukaryotic glutamine synthetase.";  
 CC Gene 129:275-278(1993).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + GLUTAMINE +  
 CC ORTHOPHOSPHATE.  
 CC -!- SUBUNIT: HOMOCCTAMER.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M96798; AAA02583.1; -  
 CC PIR: JN0716; JN0716.  
 CC HSSP: P06201; 2LGS.  
 CC PROSITE: PS00180; GLNA\_1; 1.  
 CC DR PROSITE: PS00181; GLNA\_ATP; 1.  
 CC DR PFAM: PF00120; gln-synt; 1.  
 CC KW Ligase.  
 CC SQ SEQUENCE 361 AA; 40768 MW; 3D8C3C507676099C CRC64;

Query Match 5.2%; Score 92; DB 1; Length 361;  
 Best Local Similarity 27.0%; Pred. No. 5.43e-01;  
 Matches 20; Conservative 22; Mismatches 26; Indels 6; Gaps 6;

Db 16 LDIPQKQAMVYV-DGTGENLRKSTRLNF-TPKSPSELPF-WNFDGSSGQAERSNS 72  
 QY 86 LD-AEQRDILYFSRNGTHPVFTASRWVYLYKPSPKHVRFSWHQDAVTVTCSLSYG 144  
 Db 73 DVLYLPVAVYRDPF 86  
 QY 145 DL-LYEVO-YRSPF 156

RESULT 15  
 ID Y290\_HAEIN STANDARD; PRT; 722 AA.  
 AC P77868;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE PROBABLE CATION-TRANSPORTING ATPASE H10290 (EC 3.6.1.-).  
 GN H10290  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN-RD / KW20;  
 CC MEDLINE; 95350830.  
 CC RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 CC Kerlavage A.R., Bult C.J., Tomb J.-P., Dougherty B.A., Merrick J.M.,  
 CC McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 CC Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 CC Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 CC Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 CC Fine L.B., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 CC Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 CC Venter J.C.;  
 CC "Whole-genome random sequencing and assembly of Haemophilus  
 CC influenzae Rd.";  
 CC RT Influenzae Rd.;  
 CC RL Science 269:496-512(1995).  
 CC [2]  
 CC REVISIONS.

CC White O., Kerlavage A.R., Fleischmann R.D.;  
 CC Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY.  
 CC (EI-E2 ATPASES).  
 CC -!- SIMILARITY: CONTAINS A COPY OF THE HEAVY-METAL-ASSOCIATED (HMA)  
 CC DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

Thu May 11 06:49:30 2000

```

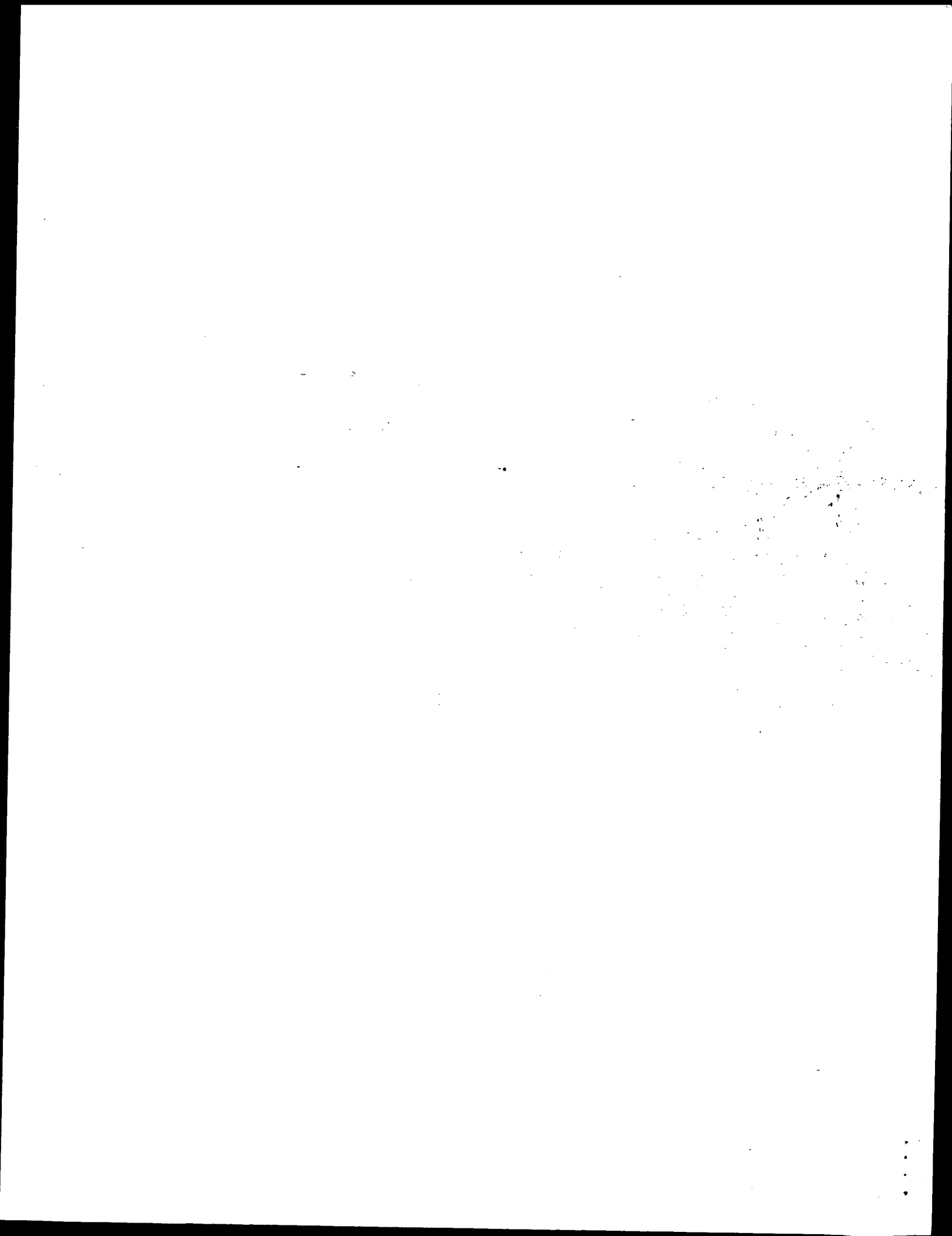
CC -----
DR EMBL; U32715; AAC21955.1; -.
DR HSP; Q04656; 2AWO.
DR TIGR; H10290; -.
DR PROSITE; PS00154; ATPASE_E1_E2; 1.
DR PROSITE; PS01047; HMA; 1.
DR PFAM; PF00122; E1-E2_ATPase; 2.
DR PFAM; PF00403; HMA; 1.
KW Hypothetical protein; Hydrolase; Transmembrane; Phosphorylation;
KW ATP-binding; Metal-binding.
FT DOMAIN 15 44 HMA.
FT METAL 20 20 POTENTIAL.
FT METAL 23 23 POTENTIAL.
FT TRANSMEM 94 114 POTENTIAL.
FT TRANSMEM 118 138 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 180 200 POTENTIAL.
FT TRANSMEM 340 360 POTENTIAL.
FT TRANSMEM 373 393 POTENTIAL.
FT TRANSMEM 523 543 POTENTIAL.
FT TRANSMEM 608 628 POTENTIAL.
FT TRANSMEM 675 695 POTENTIAL.
FT MOD_RES 697 717 POTENTIAL.
FT MOD_RES 722 722 PHOSPHORYLATION (PROBABLE).
SQ SEQUENCE 722 AA; 78070 MW; E4FF0BA5642EDCDD CRC64;

Query Match 5.2%; Score 93; DB 1; Length 722;
Best Local Similarity 30.4%; Pred. No. 3.87e-01;
Matches 17; Conservative 12; Mismatches 22; Indels 5; Gaps 4;

Db 93 RLIVLWIIPELIG--M-LGMIGGSHNMLPPIWOFALASIVOLWLAIPFYRGAI 145
QY 3 RLIVLWGAAY-FLLGGWALGOGGAEGVOIQIY-FNLEIVQVTWNASKYSTNL 56

```

Search completed: Wed May 10 11:21:08 2000  
Job time : 84 secs.



\*\*\*\*\*  
 M P S R C H  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 11:21:26 2000; Maspar time 249.04 Seconds  
 Tabular output not generated. 64,312 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (1-231) from US09376430A.pep (3 of 25)  
 Perfect Score: 1773  
 Sequence: 1 MGRVLVLLGAAGVFLGGWMA.....GEIRDACAETPTPKPKLSK 231

Scoring table: PAM 150  
 Gap 11

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: sprembl12  
 1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human  
 5:sp.invertebrate 6:sp.mammal 7:sp.mhc 8:sp.orcanelle  
 9:sp.phage 10:sp.plant 11:sp.rodent 12:sp.unclassified  
 13:sp.vertibrate 14:sp.virus

Statistics: Mean 43.890; Variance 72.209; scale 0.608

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	132	7.4	383	11	IL-13 RECEPTOR ALPHA 2	2.66e-07
2	109	6.1	435	2	PUTATIVE ASPARTATE AMI	2.53e-03
3	109	6.1	435	2	ASPARTATE AMINOTRANSF	2.53e-03
4	104	5.9	430	1	430AA LONG HYPOTHETICA	1.63e-02
5	100	5.6	890	11	INTERLEUKIN-5 RECEPT	6.95e-02
6	99	5.6	4436	1	4436AA LONG HYPOTHETIC	9.93e-02
7	99	5.6	6048	5	UNC-22 PROTEIN.	9.93e-02
8	99	5.6	6831	5	UNC-22 PROTEIN.	9.93e-02
9	99	5.6	7160	5	UNC-22 PROTEIN.	9.93e-02
10	97	5.5	256	11	CYTOKINE RECEPTOR COM	2.01e-01
11	97	5.5	311	2	SIMILAR TO TRIMETHYLAM	2.01e-01
12	98	5.5	1409	2	ALKALINE PHOSPHATASE.	1.41e-01
13	95	5.4	291	2	TYPE IV PREPILIN PEPTI	4.03e-01
14	95	5.4	315	11	MOR 5'BETA3.	4.03e-01
15	95	5.4	638	6	GROWTH HORMONE RECEPT	4.03e-01
16	95	5.4	673	14	PUTATIVE RNA DEPENDENT	4.03e-01
17	96	5.4	896	11	INTERLEUKIN-3 RECEPT	2.85e-01
18	95	5.4	987	5	SIMILAR TO IMMUNOGLOBU	4.03e-01
19	96	5.4	1082	2	PUTATIVE ARABINOSYL TR	2.85e-01
20	96	5.4	1082	2	PUTATIVE ARABINOSYL TR	2.85e-01

21	94	5.3	285	2	P73514	5.69e-01
22	94	5.3	407	5	Q9XU71	5.69e-01
23	94	5.3	919	3	Q21477	5.69e-01
24	94	5.3	981	3	Q74419	5.69e-01
25	94	5.3	1122	5	O61460	5.69e-01
26	94	5.3	1879	2	P72938	5.69e-01
27	94	5.3	26926	4	Q10466	5.69e-01
28	93	5.2	173	5	Q18307	8.00e-01
29	92	5.2	202	14	P90279	1.12e+00
30	92	5.2	349	11	Q61190	1.12e+00
31	92	5.2	684	13	Q90544	1.12e+00
32	92	5.2	1395	5	O44924	1.12e+00
33	91	5.1	202	14	P88435	2.19e+00
34	90	5.1	204	14	Q74920	2.19e+00
35	90	5.1	204	14	Q74919	2.19e+00
36	90	5.1	204	14	Q74921	2.19e+00
37	90	5.1	256	10	O65756	2.19e+00
38	91	5.1	291	2	O54483	1.57e+00
39	90	5.1	344	5	P91017	2.19e+00
40	90	5.1	376	1	O9V8P6	2.19e+00
41	90	5.1	427	4	O95646	2.19e+00
42	90	5.1	935	2	O31000	2.19e+00
43	90	5.1	1098	2	P72030	2.19e+00
44	90	5.1	1203	11	Q61984	2.19e+00
45	89	5.0	200	14	O41180	3.04e+00

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	383	AA.
ID	O88786				
AC	O88786;				
DT	01-NOV-1998 (Tremblrel. 08, Created)				
DT	01-NOV-1998 (Tremblrel. 08, Last sequence update)				
DT	01-NOV-1998 (Tremblrel. 08, Last annotation update)				
DE	IL-13 RECEPTOR ALPHA 2.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C3H/HEJ; TISSUE=THYMUS;				
RA	MEDLINE: J9391042.				
RA	DONALDSON D.D., WHITTERS M.J., FITZ L., NEBEN T.Y., FINNERTY H.,				
RA	HENDERSON S.L., O'HARA R.M. JR., BEIER D.R., TURNER K.J., WOOD C.R.,				
RA	COLLINS M.				
RT	"The murine IL-13 receptor alpha 2: molecular cloning,				
RT	characterization, and comparison with murine IL-13 receptor alpha 1."				
RL	J. Immunol. 161:2317-2324(1998).				
DR	EMBL: U65747; AAC33240.1;				
SQ	SEQUENCE 383 AA; 44483 MW; 5EAEF3E3 CRC32;				

Query Match	7.4%	Score 132;	DB 11;	Length 383;
Best Local Similarity	30.0%	Pred. No. 2.66e-07;		
Matches	27;	Conservative 25;	Mismatches 31;	Indels 7; Gaps 6;

Db	123	GISEGSLTKIQDMKCIYYNQYLVCWKPKGVSDTYMTFFVYEGLDHAL-QCADDY 181
QY	20	ALGGGAAEG-VQ-IQIIVFNLETQVQTNASK--YSRNLTFHYRFG-DEAYDOCTNY 74
Db	182	LQHDKNVGGKLSNDSYKDFFCVNGS 211
QY	75	LLEGHTSGCLLDAEQRDDIL-YFSIRNGT 103

RESULT	2	PRELIMINARY;	PRT;	435	AA.
ID	O06685				
AC	O06685;				
DT	01-JUL-1997 (Tremblrel. 04, Created)				
DT	01-JUL-1997 (Tremblrel. 04, Last sequence update)				
DT	01-NOV-1998 (Tremblrel. 08, Last annotation update)				
DE	PUTATIVE ASPARTATE AMINOTRANSFERASE TPAAT.				
GN	TPAAT.				

[illegible]



US-09-876-430-2-03.isrpt

Thu May 11 06:49:31 2000

```

ID O58659 PRELIMINARY; PRT; 4436 AA.
AC O58659; 1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE 4436AA LONG HYPOTHETICAL PROTEIN.
GN PH0954.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE; 98344137.
RX KAWABAYASHI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,
RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOIYAMA A., NAGAI Y.,
RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,
RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,
RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,
RA KIKUCHI H.
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
DR EMBL; AP000004; BAA30051.1; -.
DR PFAM; PF000041; fn3; 1.
DR PRINIS; PR00801; PKD; 2.
DR PRINIS; PR00801; PKD; 2.
SQ SEQUENCE 4436 AA; 497015 MW; 2D9608CA CRC32;

Query Match 5.6%; Score 99; DB 1; Length 4436;
Best Local Similarity 26.8%; Pred. No. 9.93e-02;
Matches 19; Conservative 18; Mismatches 29; Indels 5; Gaps 5;

Db 4115 NMTYNTREIKVPAT-D-E-SGIANTATINGSLSLEKVNWTGIRVLDGKVELNV 4171
QY 140 DLSYGDLLYEVOYRSPDTEWQSKQENTCNVTIEGLDAEKCYSEFW-RVKAMEDVYGPDT 198

Db 4172 FASDKWGNVC 4182
QY 199 YPSD-WSEVTC 208

RESULT 7
ID Q23020 PRELIMINARY; PRT; 6048 AA.
AC Q23020; Q27232;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE TWITCHIN.
GN UNC-22 OR ZK617.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae.
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE; 93367654.
RA BENIAN G.M., L'HERNAULT S.W., MORRIS M.E.;
RT "Additional sequence complexity in the muscle gene, unc-22, and its
RT encoded protein, twitchin, of Caenorhabditis elegans."
RT Genetics 134:1097-1104(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA HARRIS B.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X15423; CAA33463.1; -.
DR EMBL; 273899; CAA98081.1; ALT_INIT.

```

```

DR EMBL; 273897; CAA98081.1; JOINED.
DR HSP; Q63450; 1A06.
DR PFAM; PF00041; fn3; 31.
DR PFAM; PF00047; ig; 13.
DR PFAM; PF00069; pkinase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR MYOSIN; Kinase.
SQ SEQUENCE 6048 AA; 668449 MW; 1977C602 CRC32;

Query Match 5.6%; Score 99; DB 5; Length 6048;
Best Local Similarity 33.9%; Pred. No. 9.93e-02;
Matches 19; Conservative 8; Mismatches 27; Indels 2; Gaps 2;

Db 916 YIVEVRDPDKWKEVKRVPDNTASISGLKEGKEYOFRRVAVNKAGP-GQPSEPE 970
QY 148 YEVOYRSPDTEW-OSKQENTCNVTIEGLDAEKCYSEFWRVKAMEDVYGPDTYPSD 202

RESULT 8
ID Q23550 PRELIMINARY; PRT; 6831 AA.
AC Q23550;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE UNC-22 PROTEIN.
GN UNC-22.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae.
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA WHITE S.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 273897; CAA98064.1; -.
DR EMBL; 273899; CAA98064.1; JOINED.
DR HSP; P02751; 1FNA.
DR PFAM; PF00041; fn3; 31.
DR PFAM; PF00047; ig; 17.
DR PFAM; PF00069; pkinase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINIS; PR00014; FNTYPEIII.
SQ SEQUENCE 6831 AA; 752579 MW; 0A66C338 CRC32;

Query Match 5.6%; Score 99; DB 5; Length 6831;
Best Local Similarity 33.9%; Pred. No. 9.93e-02;
Matches 19; Conservative 8; Mismatches 27; Indels 2; Gaps 2;

Db 1699 YIVEVRDPDKWKEVKRVPDNTASISGLKEGKEYOFRRVAVNKAGP-GQPSEPE 1753
QY 148 YEVOYRSPDTEW-OSKQENTCNVTIEGLDAEKCYSEFWRVKAMEDVYGPDTYPSD 202

RESULT 9
ID Q23551 PRELIMINARY; PRT; 7160 AA.
AC Q23551;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE ZK617.1B PROTEIN.
GN ZK617.1B.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae.
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA HARRIS B.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA WILSON R., AINSICOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,
RA BONFIELD J., BURTON J., CONNELL M., COOPER J., COOPER J., COOPER J.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA

```

RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFFEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans";  
 RL Nature 368:32-38(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA WHITE S.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z73899; CAA98082.1; -.  
 DR EMBL: Z73897; CAA98082.1; JOINED.  
 DR EMBL: Z73897; CAA98085.1; -.  
 DR EMBL: Z73899; CAA98065.1; JOINED.  
 DR HSSP: P02751; LFNA.  
 DR PFAM: PF00041; fn3; 31.  
 DR PFAM: PF00047; ig; 17.  
 DR PFAM: PF00069; pkinase; 1.  
 DR PRINTS: PRO0014; ENYPEI1.  
 SQ SEQUENCE 7160 AA; 789211 MW; EDD567FE CRC32;

Query Match 5.6%; Score 99; DB 5; Length 7160;  
 Best Local Similarity 33.9%; Pred. No. 9.93e-02;  
 Matches 19; Conservative 8; Mismatches 27; Indels 2; Gaps 2;  
 Db 2028 YIVEVRDPTKEWKEVKRVPDTNATISGLKEGKEVQFVRVAVNKAQP-GQPSSE 2082  
 QY 148 YEVQSRSPFDIEW-QSKQENTCNVTIEGLDAEKCYSFWRVVRKAMEDVYGPDTYPSD 202

RESULT 10  
 ID Q63968 PRELIMINARY; PRT; 256 AA.  
 AC Q63968;  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DE CYTOKINE RECEPTOR COMPLEX COMMON BETA CHAIN H BETA C (FRAGMENT).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 94235843.  
 RA D'ANDREA R., RAYNER J., MORETTI P., LOPEZ A., GOODALL G.J.,  
 RA GONDA T.J., VADAS M.;  
 RT "A mutation of the common receptor subunit for interleukin-3 (IL-3),  
 RT to ligand independence colony-stimulating factor, and IL-5 that leads  
 RT to ligand independence and tumorigenicity.";  
 RL Blood 83:2802-2808(1994).  
 DR EMBL: S70302; AAB31055.1; -.  
 DR HSSP: P40189; 1BQI.  
 DR PFAM: PF00041; fn3; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 256 AA; 29206 MW; 5242B76B CRC32;

Query Match 5.5%; Score 97; DB 11; Length 256;  
 Best Local Similarity 22.6%; Pred. No. 2.01e-01;  
 Matches 14; Conservative 16; Mismatches 29; Indels 3; Gaps 3;  
 Db 80 RTGYNGINSESEAH-SMALPALEPSTRYARVRVTSRTGYN-GIW-SEWSEARSWDTE 136  
 QY 153 RSPFDTEWQSKQENTCNVTIEGLDAEKCYSFWRVVRKAMEDVYGPDTYPSDSEVTCWQ 212  
 Db 137 SV 138  
 QY 213 EI 214

RESULT 11  
 ID Q48962 PRELIMINARY; PRT; 311 AA.

AC Q48962;  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DE SIMILAR TO TRIMETHYLAMINE DH (FRAGMENT).  
 OS Mycoplasma capricolum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC capricolum group.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 27343(KID);  
 RX MEDLINE: 96059641.  
 RA BORK P., OUZOUNIS C., CASARI G., SCHNEIDER R., SANDER C., DOLAN M.,  
 RA GILBERT W., GILLET P.M.;  
 RT "Exploring the Mycoplasma capricolum genome: a minimal cell reveals  
 RT its physiology";  
 RL Mol. Microbiol. 16:955-967(1995).  
 DR EMBL: Z33015; CAA83700.1; -.  
 DR PFAM: PF00724; Oxidored\_FMN; 1.  
 FT NON\_TER 311  
 SQ SEQUENCE 311 AA; 35768 MW; 6CC72E66 CRC32;

Query Match 5.5%; Score 97; DB 2; Length 311;  
 Best Local Similarity 27.9%; Pred. No. 2.01e-01;  
 Matches 12; Conservative 11; Mismatches 19; Indels 1; Gaps 1;  
 Db 202 FCLEVVKAIREVIDKYAPKNEIFGFRAPEETGYDILGYTTIED 244  
 QY 37 FNLETVQVTWNA-SKYSRTNLTFFHFRNGDAYDQCTNYLLOE 78

RESULT 12  
 ID P72939 PRELIMINARY; PRT; 1409 AA.  
 AC P72939;  
 DT 01-FEB-1997 (TEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)  
 DE ALKALINE PHOSPHATASE.  
 OS Synechocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCC6803;  
 RA TABATA S.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PCC6803;  
 RX MEDLINE: 97061201.  
 RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,  
 RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAKI M., SASAKI T.,  
 RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,  
 RA SHIMO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,  
 RA TABATA S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 DR EMBL: D90902; BAA16956.1; -.  
 DR PFAM: PF01009; 5\_nucleotidase; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 1409 AA; 149390 MW; 7597586F CRC32;

Query Match 5.5%; Score 98; DB 2; Length 1409;  
 Best Local Similarity 41.7%; Pred. No. 1.41e-01;  
 Matches 15; Conservative 9; Mismatches 9; Indels 3; Gaps 3;  
 Db 1336 LGNWOALA-AETVQGVN-QVLWQNLDTNIGVWNS 1369  
 QY 15 LGGWMLGOGGAEGVQIQIYYENLEIVQV-TWNAS 49

RESULT 13  
 ID O68964 PRELIMINARY; PRT; 291 AA.

C		O68964;	
JT		01-AUG-1998 (TriEMBLrel. 07, Created)	
T		01-AUG-1998 (TriEMBLrel. 07, Last sequence update)	
DJ		01-NOV-1999 (TriEMBLrel. 12, Last annotation update)	
OT		TYPE IV PREPILIN PEPTIDASE TAPD.	
S		TAPD.	
S		Aeromonas salmonicida.	
OC		Bacteria; Proteobacteria; gamma subdivision; Aeromonas group;	
CC		Aeromonas.	
N	[1]	SEQUENCE FROM N.A.	
RN		STRAIN-A450;	
RP		PEPE C.M., STROM M.S.;	
RC		Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.	
RA		EMBL: AF059249; AAC23569.1;	
RRL		PFAM: PF01478; Peptidase_C20; 1.	
DR		PRINTS: PR00864; PREPILNPBASE.	
CSQ		SEQUENCE 291 AA; 32500 MW; 1A29B48C CRC32:	
		Query Match 5.4%; Score 95; DB 2; Length 291;	
		Best Local Similarity 54.5%; Pred. No. 4.03e-01;	
	Matches	12; Conservative 6; Mismatches 3; Indels 1; Gaps 1;	
Ddb	164	LPLLWGGLLFNLGGYVPLGDA 185	
Qy	4	LVLWGAARVF-LLGWMALQGQ 24	

```

RESULT 14
AC Q9WVN6 PRELIMINARY; PRT; 315 AA.
ID Q9WVN6
AD Q9WVN6;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE MOR 5'BETA3.
DI OS Mus musculus (Mouse).
OS Mus musculus; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eukaryota; Metazoa; Chordata; Scleroognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-129;
RC MEDLINE; 99055560.
RX BENDER M.A., REIK A., CLOSE J., TELLING A., EPNER E., FIERING S.,
RA HARDISON R., GROUDINE M.;
RT "description and targeted deletion of 5' hypersensitive site 5 and 6
of the mouse beta-globin locus control region.";
RL Blood 92:4394-4403(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-129;
RC MEDLINE; 99238494.
RX BULGER M., VON DOORNICK J.H., SAITOH N., TELLING A., FARRELL C.,
RA BENDER M.A., FEISENFELD G., AXEL R., GROUDINE M.;
RT "Conservation of sequence and structure flanking the mouse and human
beta-globin loci: the beta-globin genes are embedded within an array
of odorant receptor genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:5129-5134(1999).
DR ENBL; AF071080; AAD28304.1;
SC SEQUENCE 315 AA, 35559 MW, DBE53BF7 CRC32;

```

	Query Match	5.4%	Score 95:	DB 11:	Length 315:
	Best Local Similarity	28.3%	Pred. No. 4.0e-01:		
	Matches 13:	Conservative	Matches 17:	Indels	Gaps 3:
159	PIIRLHWFYPC-RSVLSGHA-FCLHQDVVKACADITF-NRLYPV	201			
105	PVFASRWVYLYLKSPSKHVRFSWHQDAVTTCSDLYSGLLYEV	150			

RESULT	15	
ID	Q9XSZ1	PRELIMINARY; PRT; 638 AA.
AC	Q9XSZ1;	
DT	01-NOV-1999	(TREMBlrel. 12, Created)

```

01-NOV-1999 (TREWBlrel. 12, Last sequence update)
01-NOV-1999 (TREWBlrel. 12, Last annotation update)
G1-GROWTH HORMONE RECEPTOR.
Papio anubis (Olive baboon).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae;
Papio.
[1]
SEQUENCE FROM N.A.
RN
RP
RZ
RT
RA
RT
RT
RT
RT
DR
KW
SQ
SEQUENCE 638 AA: 71407 MW; 2EC386D7 CRC32;

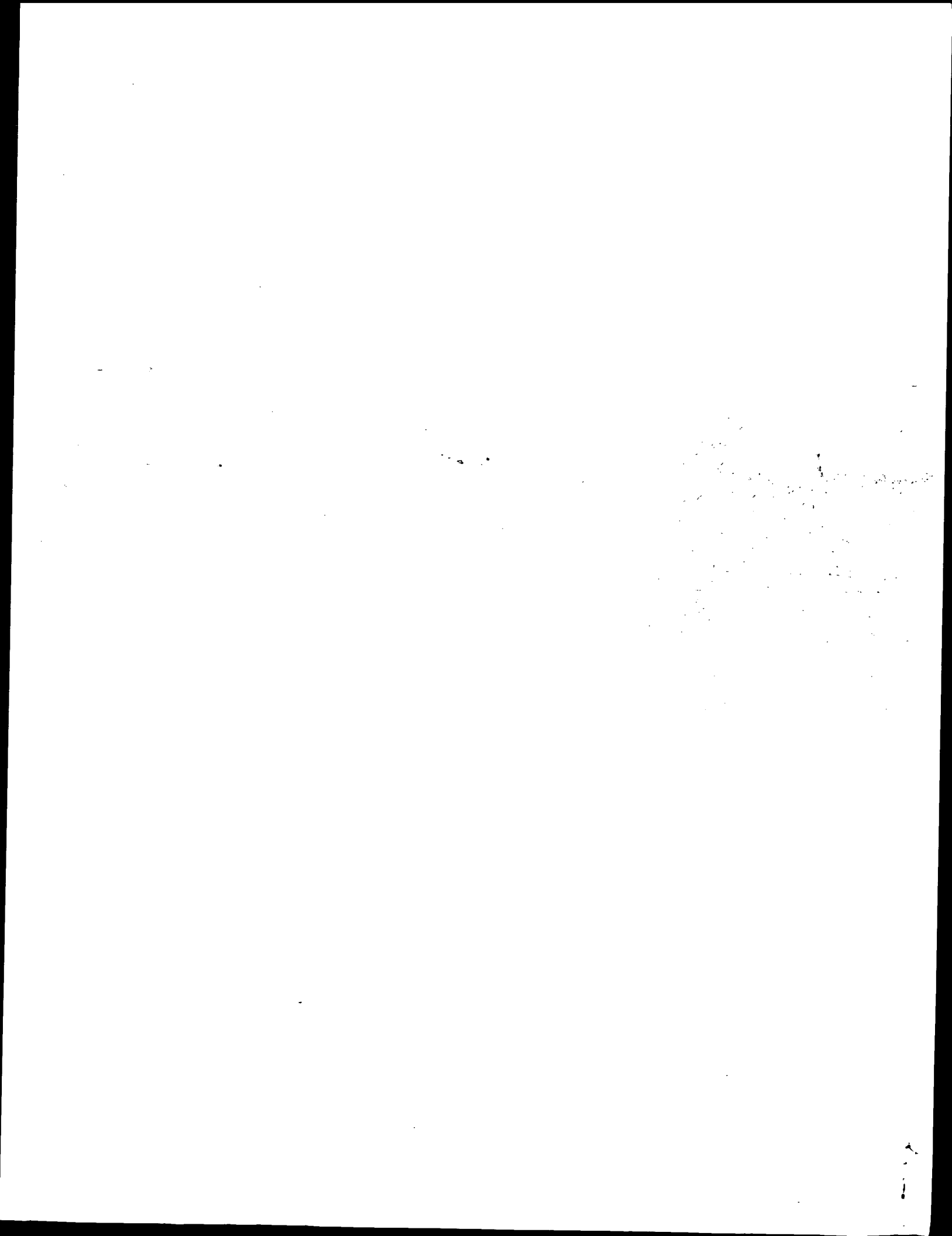
Query Match          5.4%; Score 95; DB 6: Length 638;
Best Local Similarity 30.8%; Pred. NO. 4.03e-01;
Matches 16; Conservative 12; Mismatches 22; Indels 2; Gaps 2;

Db 190 LEVELOYKEVNETKWKMPDILSTSPVYSLKVDKEYEVRYSKRNSNGY 241
| | | | | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 146 LLXEYQYRSPDTEWOSKOEN-TCNVITIEGLDAEKCYSFVVRKAMEDV-YG 195

Search completed: wed May 10 11:25:46 2000
Job time : 260 secs.

```

Search completed: Wed May 10 11:25:46 2000  
Job time : 260 secs.



Thu May 11 06:49:31 2000

\*\*\*\*\*  
M P S R L H  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:35:43 2000; MasPar time 9.00 Seconds  
Tabular output not generated. 549.891 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (23-231) from US09376430A.ppt (4 of 25)  
Perfect Score: 1586  
Sequence: 1 QGGAGVGQIIYFNLETV.....GEIRDCAETPPPKPLSK 209

Scoring table: PAM 150  
Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq35  
1:geneseqp

Statistics: Mean 31.996; Variance 121.649; scale 0.263

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	200	12.6	482	1 W31646	Human cytokine receptor	1.54e-09
2	195	12.3	230	1 R82934	Interleukin 4 component	4.76e-09
3	195	12.3	230	1 R47151	IL-2 receptor gamma ch	4.76e-09
4	195	12.3	252	1 R47150	IL-2 receptor gamma ch	4.76e-09
5	195	12.3	347	1 R47149	IL-2 receptor gamma ch	4.76e-09
6	195	12.3	369	1 R47148	IL-2 receptor gamma ch	4.76e-09
7	188	11.9	369	1 R59094	Murine IL-2R gamma.	2.30e-08
8	130	8.2	383	1 W32594	Murine IL-13 binding c	6.80e-03
9	115	7.3	576	1 R78613	Expression vector pME1	1.46e-01
10	115	7.3	596	1 R78616	Expression vector pME1	1.46e-01
11	115	7.3	600	1 R92526	Fas antigen #1.	1.46e-01
12	115	7.3	600	1 R78610	Expression vector pME1	1.46e-01
13	115	7.3	878	1 R78608	Murine IL-3 receptor b	1.46e-01
14	112	7.1	592	1 R32527	Fas antigen #2.	2.67e-01
15	113	7.0	237	1 R22229	Sequence of beta-chain	2.19e-01
16	111	7.0	269	1 R05045	Truncated human growth	3.26e-01
17	111	6.9	237	1 R24274	Soluble human growth h	3.98e-01
18	110	6.9	878	1 R32529	Truncated human growth	4.85e-01
19	109	6.9	572	1 Y04954	Fas sequence from AIC2	1.06e-00
20	105	6.6	110	1 R36390	Myobacterium species	2.81e-00
21	100	6.3	112	1 R36391	Extracellular hGHR-CD	2.81e-00
22	100	6.3	112	1 R36391	Extracellular hGHR-CD	2.81e-00
23	100	6.3	246	1 R56389	Human growth hormone r	2.81e-00

24	100	6.3	249	1 R06867	Hormone binding region	2.81e+00
25	100	6.3	269	1 W10426	Human somatogenic rece	2.81e+00
26	100	6.3	637	1 P92108	Human growth hormone r	2.81e+00
27	100	6.3	638	1 W33394	Human growth hormone r	2.81e+00
28	100	6.3	638	1 P81326	Human growth hormone r	2.81e+00
29	94	5.9	269	1 W82802	Human soluble somatoge	8.79e+00
30	93	5.9	315	1 W56261	Mature interleukin-13	1.06e+01
31	93	5.9	359	1 W56260	Construct containing m	1.06e+01
32	93	5.9	380	1 W41502	Human cytokine/peptide	1.06e+01
33	93	5.9	380	1 W41502	Human interleukin-13 b	1.06e+01
34	93	5.9	380	1 W41520	Human HR-1 receptor.	1.06e+01
35	93	5.9	380	1 W33603	Homo sapiens HR-1 rece	1.06e+01
36	93	5.9	380	1 W35295	Human IL-13 binding ch	1.06e+01
37	93	5.9	380	1 W36613	Human zcytor2 cytokine	1.06e+01
38	94	5.9	638	1 W33395	Rabbit growth hormone	8.79e+00
39	94	5.9	638	1 P81327	Rabbit growth hormone	8.79e+00
40	94	5.9	638	1 P92107	Rabbit growth hormone	8.79e+00
41	91	5.7	380	1 W56614	Human zcytor2 cytokine	1.85e+01
42	90	5.7	426	1 W09822	Human interleukin-12 r	1.85e+01
43	89	5.7	427	1 W24973	Human interleukin-13 a	2.23e+01
44	88	5.6	446	1 R06448	C. jejuni hippuricase	2.23e+01
45	88	5.5	372	1 W36616	Celebus macaque zcytor	2.68e+01

ALIGNMENTS

RESULT 1  
ID W31646 standard; Protein: 482 AA.

AC	W31646;					
DT	21-MAY-1998 (first entry)					
DE	Human cytokine receptor gc chain-Ig fusion protein.					
KW	Cytokine receptor; gamma common chain; gc chain; human;					
KW	blocking agent; monoclonal antibody; CP.B8; immunological disease;					
KW	myasthenia gravis; rheumatoid arthritis; lupus; multiple sclerosis;					
KW	insulin-dependent diabetes; inflammatory bowel disease;					
KW	sympathetic ophthalmia; uveitis; allergy; asthma; infection;					
KW	graft versus host disease; psoriasis; immunosuppressive; therapy.					
OS	Chimeric - Homo sapiens					
FH	Key					
FT	Protein	1..254				
FT	Protein	1..254				
FT	Protein	255..482				
FT	Protein	255..482				
FT	Region	255..284				
FT	Region	255..284				
FT	Domain	264..482				
FT	Domain	264..482				
PN	WO9743416-A1.					
PD	20-NOV-1997.					
PR	09-MAY-1997; U07870.					
PR	10-MAY-1996; US-017466.					
PI	(BIOJ) BIOGEN INC.					
PI	Benjamin CD, Burkly LC, Hession C, Whitty A;					
DR	WPI: 98-008885/01.					
DR	N-PSDB: T97439.					
PT	Blocking agents of the gamma common chain of cytokine receptors -					
PT	treatment of immunological diseases					
PS	Example 1; Page 79-80; 111pp; English.					
CC	This polypeptide comprises a fusion between the N-terminal 254					
CC	amino acids of the human mature cytokine receptor gamma common (gc)					
CC	chain and the hinge region and CH2 and CH3 constant domains of					
CC	human IgG1. The fusion was expressed from clone pL8001 (see					
CC	W97439) in COS-7 cells, and used to generate murine anti-human gc					
CC	specific monoclonal antibodies (MAbs), including CP.B8 produced by					
CC	hybridoma ATCC HB 12107. The invention provides compositions and					
CC	methods for inhibiting cytokine signalling using gc chain blocking					
CC	agents for the treatment of immunological diseases such as					
CC	myasthenia gravis, rheumatoid arthritis, lupus, multiple sclerosis,					
CC	insulin-dependent diabetes, inflammatory bowel disease, syphathetic					
CC	ophthalmia, uveitis, allergy, asthma, parasitic infection, graft					
CC	vs. host disease or psoriasis. A preferred gc blocking agent is					
CC	Mab CP.B8 or its Fab fragment (see also W31647-48).					

SQ Sequence 482 AA;  
 Query Match 12.6%; Score 200; DB 1; Length 482;  
 Best Local Similarity 25.7%; Pred. No. 1.54e-09;  
 Matches 53; Conservative 53; Mismatches 88; Indels 12; Gaps 12;  
 Db 59 EVQCFVFNVEYMNCTWNSSEPPQNTLTHYKNSDNDKVKCSHYLFSEITSGCOLQ 118  
 QY 31 QIQIYFNLETVQVTWNAKYSR-TNLTFFHYRF-NGD-EAYDQCTNVLQEGHTSGCLLD 87  
 Db 119 KKEIHLVQTFVYVQLQDPRPRQATQMLKQNLVWPAPENLTLLHKLSESOLELNNWNR 178  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYIL-KPSPKHYRF-SWHQDAVTTCSDLS 142  
 Db 179 LNHCHLHVQYRTDWDHSWTEQSDYDRHKFSLPSVDGQKRYMFRVRSR-FNPLCGSAQHW 237  
 QY 143 YGD-LLYEVQYRSPFDTEW-QSKQENTCNVTIEGLDAEKCYSFVWVRKAMEDVYGPDTYP 200  
 Db 238 SEWSHPITHGNSNTSKENVDKTHTCPP 263  
 QY 201 SDWSEVTCWQGEIRDAETPT-PP 225

RESULT 2  
 ID R82934 standard; Protein; 230 AA.  
 AC R82934;  
 DE Interleukin-4 component common to the IL-2 receptor gamma chain.  
 KW Interleukin-4; IL-4; gamma chain component; immunosuppressants;  
 KW anti-allergy agent; signal transduction inhibitor; autoimmune;  
 KW disease; anti-inflammatory; anaphylactic shock; bronchial asthma;  
 KW Interleukin-2; IL-2; atopic dermatitis; urticaria.  
 OS Homo sapiens.  
 PN J07149662-A.  
 PD 13-JUN-1995.  
 PF 07-SEP-1994; 213706.  
 PR 08-SEP-1993; JP-223574.  
 PA (AJIN) AJINOMOTO KK.  
 PA (SUGA) SUGAMURA K.  
 DR WPI; 95-243601/32.  
 DR N-PSDB; T04952.  
 PT Novel interleukin-4 receptor monoclonal antibodies inhibit signal  
 PT transmission - useful as immunosuppressants and anti-allergy agents.  
 PS Example 1; Page 9; lipp; Japanese.  
 CC T04952 encodes R82934 a component of the IL-4 receptor common to  
 CC the IL-2 receptor gamma chain molecule, which was used to generate  
 CC anti-IL-4 receptor monoclonal antibodies (mAbs). The mAbs (IL-4  
 CC signal transduction inhibitors) can be used as immunosuppressants  
 CC and anti-allergy agents, for the treatment of autoimmune and chronic  
 CC inflammatory diseases, e.g. anaphylactic shock, bronchial asthma,  
 CC atopic dermatitis and urticaria.  
 SQ Sequence 230 AA;

Query Match 12.3%; Score 195; DB 1; Length 230;  
 Best Local Similarity 25.9%; Pred. No. 4.76e-09;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 38 EVQCFVFNVEYMNCTWNSSEPPQNTLTHYKNSDNDKVKCSHYLFSEITSGCOLQ 97  
 QY 31 QIQIYFNLETVQVTWNAKYSR-TNLTFFHYRF-NGD-EAYDQCTNVLQEGHTSGCLLD 87  
 Db 98 KKEIHLVQTFVYVQLQDPRPRQATQMLKQNLVWPAPENLTLLHKLSESOLELNNWNR 157  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYIL-KPSPKHYRF-SWHQDAVTTCSDLS 142  
 Db 158 LNHCHLHVQYRTDWDHSWTEQSDYDRHKFSLPSVDGQKRYMFRVRSR-FNPLCGSAQHW 216  
 QY 143 YGD-LLYEVQYRSPFDTEW-QSKQENTCNVTIEGLDAEKCYSFVWVRKAMEDVYGPDTYP 200  
 Db 217 SEWSHPITHW 225  
 QY 201 SDWSEVTCW 209

RESULT 3  
 ID R47151 standard; Protein; 230 AA.  
 AC R47151;  
 DT 13-JUN-1994 (first entry)  
 DE IL-2 receptor gamma chain.  
 KW Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;  
 KW rheumatoid arthritis; transplant rejection; primer;  
 KW polymerase chain reaction; PCR; amplification.  
 OS Homo sapiens.  
 PN EP-578932-A.  
 PD 19-JAN-1994.  
 PF 22-APR-1993; 106561.  
 PR 23-APR-1992; JP-104947.  
 PA (AJIN) AJINOMOTO KK.  
 PA (SUGA) SUGAMURA K.  
 PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
 PI Suzuki M, Takeshita T;  
 DR WPI; 94-017546/03.  
 DR P-PSDB; Q54831.  
 PT DNA and protein sequences of IL-2 gamma chain - useful as immune  
 PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
 PT transplant rejection  
 PS Disclosure; Page 22-23, 35-36; 50pp; English.  
 CC The human IL-2 receptor gamma chain preform (R47148), including the  
 CC signal peptide, is encoded by the sequence given in Q54828. The  
 CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
 CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
 CC while a soluble form suitable for expression in prokaryotes (R47151)  
 CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
 CC terminal sequence of IL-2 receptor gamma chain, and are used to  
 CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
 CC are used to obtain the protein given in R47151.  
 SQ Sequence 230 AA;

Query Match 12.3%; Score 195; DB 1; Length 230;  
 Best Local Similarity 25.9%; Pred. No. 4.76e-09;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 37 EVQCFVFNVEYMNCTWNSSEPPQNTLTHYKNSDNDKVKCSHYLFSEITSGCOLQ 96  
 QY 31 QIQIYFNLETVQVTWNAKYSR-TNLTFFHYRF-NGD-EAYDQCTNVLQEGHTSGCLLD 87  
 Db 97 KKEIHLVQTFVYVQLQDPRPRQATQMLKQNLVWPAPENLTLLHKLSESOLELNNWNR 156  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-VYIL-KPSPKHYRF-SWHQDAVTTCSDLS 142  
 Db 157 LNHCHLHVQYRTDWDHSWTEQSDYDRHKFSLPSVDGQKRYMFRVRSR-FNPLCGSAQHW 215  
 QY 143 YGD-LLYEVQYRSPFDTEW-QSKQENTCNVTIEGLDAEKCYSFVWVRKAMEDVYGPDTYP 200  
 Db 216 SEWSHPITHW 224  
 QY 201 SDWSEVTCW 209

RESULT 4  
 ID R47150 standard; Protein; 252 AA.  
 AC R47150;  
 DT 13-JUN-1994 (first entry)  
 DE IL-2 receptor gamma chain.  
 KW Interleukin-2 receptor gamma chain.  
 KW rheumatoid arthritis; transplant rejection; primer;  
 KW polymerase chain reaction; PCR; amplification; ss.  
 OS Homo sapiens.  
 PN EP-578932-A.  
 PD 19-JAN-1994.  
 PF 22-APR-1993; 106561.  
 PR 23-APR-1992; JP-104947.  
 PA (AJIN) AJINOMOTO KK.

US-09-376-430-2-04.rag

Thu May 11 06:49:31 2000

PA (SUGA/) SUGAMURA K. Nakamura M, Shimamura T, Sugamura K;  
PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
PI Suzuki M, Takeshita T;  
DR WPI: 94-017546/03.  
DR P-PSDB: Q54830.  
DR DNA and protein sequences of IL-2 gamma chain - useful as immune  
PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
PT transplant rejection  
PS Disclosure: Page 21-22, 34-35; 50pp; English.  
CC The human IL-2 receptor gamma chain preform (R47148), including the  
CC signal peptide, is encoded by the sequence given in Q54828. The  
CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
CC while a soluble form suitable for expression in prokaryotes (R47151)  
CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
CC terminal sequence of IL-2 receptor gamma chain, and are used to  
CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
CC are used to obtain the protein given in R47151.  
SQ Sequence 252 AA;

Query Match 12.3%; Score 195; DB 1; Length 252;  
Best Local Similarity 25.9%; Pred. No. 4.76e-09;  
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
Db 59 EVQCFVFNVEYVNCNWSSEPPQPTNLTHYWKNSDNKVKQKSHYLFSEITSGCOLQ 118  
QY 31 QIQIIFNLETVQVWTWNAKYSR-TNLTFHYRF-NGD-EAYDOCTNLLQEGHTSGCLLD 87  
Db 119 KKEHLHYQTFVVOLODPREPRQATOMLKLONLVPWAPENLTLHLKLSQLELNNRNF 178  
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-YYIL-KPSPKHVRF-SWHQDAVTVTCSDL 142  
Db 179 LNHCLHLVQYRTDWDHSWTSQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 237  
QY 143 YGD-LLYEVQYRSFPDTEM-QSKQENTCNVTIEGLDAEKCSFWVRKAMEDVYGPDTYP 200  
Db 238 SEWSPHWH 246  
QY 201 SDWSEVTCW 209

RESULT 5  
ID R47148 standard; Protein: 347 AA.  
AC R47149; 1994 (first entry)  
DT 13-JUN-1994  
DE IL-2 receptor gamma chain.  
KW Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;  
KW rheumatoid arthritis; transplant rejection; primer; PCR;  
KW polymerase chain reaction; PCR; amplification.  
OS Homo sapiens.  
PN EP-578932-A.  
PD 19-JAN-1994.  
PF 22-APR-1993; 106561.  
PR 23-APR-1992; JP-104947.  
PA (AJIN ) AJINOMOTO KK.  
PA (SUGA/) SUGAMURA K.  
PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
PI Suzuki M, Takeshita T;  
DR WPI: 94-017546/03.  
DR P-PSDB: Q54829.  
DR DNA and protein sequences of IL-2 gamma chain - useful as immune  
PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
PT transplant rejection  
PS Claim 4; Page 41; 50pp; English.  
CC The human IL-2 receptor gamma chain preform (R47148), including the  
CC signal peptide, is encoded by the sequence given in Q54828. The  
CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
CC while a soluble form suitable for expression in prokaryotes (R47151)  
CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
CC terminal sequence of IL-2 receptor gamma chain, and are used to  
CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
CC are used to obtain the protein given in R47151.

SQ Sequence 347 AA;  
Query Match 12.3%; Score 195; DB 1; Length 347;  
Best Local Similarity 25.9%; Pred. No. 4.76e-09;  
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
Db 37 EVQCFVFNVEYVNCNWSSEPPQPTNLTHYWKNSDNKVKQKSHYLFSEITSGCOLQ 96  
QY 31 QIQIIFNLETVQVWTWNAKYSR-TNLTFHYRF-NGD-EAYDOCTNLLQEGHTSGCLLD 87  
Db 97 KKEHLHYQTFVVOLODPREPRQATOMLKLONLVPWAPENLTLHLKLSQLELNNRNF 156  
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-YYIL-KPSPKHVRF-SWHQDAVTVTCSDL 142  
Db 157 LNHCLHLVQYRTDWDHSWTSQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 215  
QY 143 YGD-LLYEVQYRSFPDTEM-QSKQENTCNVTIEGLDAEKCSFWVRKAMEDVYGPDTYP 200  
Db 216 SEWSPHWH 224  
QY 201 SDWSEVTCW 209

RESULT 6  
ID R47148 standard; Protein: 369 AA.  
AC R47148;  
DT 13-JUN-1994 (first entry)  
DE IL-2 receptor gamma chain.  
KW Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;  
KW rheumatoid arthritis; transplant rejection; primer; PCR;  
KW polymerase chain reaction; PCR; amplification; ss.  
OS Homo sapiens.  
PN EP-578932-A.  
PD 19-JAN-1994.  
PF 22-APR-1993; 106561.  
PR 23-APR-1992; JP-104947.  
PA (AJIN ) AJINOMOTO KK.  
PA (SUGA/) SUGAMURA K.  
PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
PI Suzuki M, Takeshita T;  
DR WPI: 94-017546/03.  
DR P-PSDB: Q54828.  
DR DNA and protein sequences of IL-2 gamma chain - useful as immune  
PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
PT transplant rejection  
PS Disclosure: Page 16-17, 29-30; 50pp; English.  
CC The human IL-2 receptor gamma chain preform (R47148), including the  
CC signal peptide, is encoded by the sequence given in Q54828. The  
CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830,  
CC while a soluble form suitable for expression in prokaryotes (R47151)  
CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
CC terminal sequence of IL-2 receptor gamma chain, and are used to  
CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
CC are used to obtain the protein given in R47151.  
SQ Sequence 369 AA;

Query Match 12.3%; Score 195; DB 1; Length 369;  
Best Local Similarity 25.9%; Pred. No. 4.76e-09;  
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
Db 59 EVQCFVFNVEYVNCNWSSEPPQPTNLTHYWKNSDNKVKQKSHYLFSEITSGCOLQ 118  
QY 31 QIQIIFNLETVQVWTWNAKYSR-TNLTFHYRF-NGD-EAYDOCTNLLQEGHTSGCLLD 87  
Db 119 KKEHLHYQTFVVOLODPREPRQATOMLKLONLVPWAPENLTLHLKLSQLELNNRNF 178  
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRW-YYIL-KPSPKHVRF-SWHQDAVTVTCSDL 142  
Db 179 LNHCLHLVQYRTDWDHSWTSQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 237

1



US-09-376-430-2-04.rag

Thu May 11 06:49:31 2000

KW extracellular; region; AIC2A; soluble membrane protein;  
 OS antibody production; diseases; treatment; prevention.  
 AC J07115988-A.  
 AC Mus musculus.  
 AC J07115988-A.  
 AC 09-MAY-1995.  
 AC 26-OCT-1993; JP-267644.  
 AC 26-OCT-1993; JP-267644.  
 AC (NIBS ) JAPAN TOBACCO INC.  
 AC WPI: 95-202847/27.  
 AC N-PSDB: Q95306.  
 AC Preparation of soluble membrane proteins - for their use in antibody  
 PT production for the treatment and prevention of related diseases  
 PS Claim 10; Pages 36-38; 51pp; Japanese.  
 CC R78613 is the protein prod. of the expression vector pME18S/murine Fas  
 CC antigen-extracellular region-AIC2A. The expression vector was used for  
 CC the prodn. of recombinant soluble membrane proteins. The proteins can  
 CC be used in antibody prodn. for the treatment and prevention of related  
 CC diseases.  
 CC Sequence 576 AA;  
 SQ

Query Match 7.3%; Score 115; DB 1; Length 576;  
 Best Local Similarity 32.4%; Pred. No. 1.46e-01;  
 Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 502 YIDHFFQVQYKKSES-WKDSKTENLGRVNSMDLPQLEPDTSYCARVRVKPISD-YD-GI 558  
 QY 143 YGDLLEYVQYRSPFTEW-QSKQENTCNV-TIE-G-LDAEKCYSEFWVRKAMEDVYGPDT 198  
 Db 559 W-SEWSNEYTW 568  
 QY 199 YPSDWSEVTCW 209

RESULT 10  
 ID R78616 standard; Protein; 596 AA.  
 AC R78616;  
 AC 20-FEB-1996 (first entry)  
 DE Expression vector pME18S/MPD-1, EXT-AIC2A protein prod.  
 DE Expression vector; pME18S/MPD-1, EXT-AIC2A;  
 KW extracellular; region; AIC2A; soluble membrane protein;  
 KW antibody production; diseases; treatment; prevention.  
 OS Mus musculus.  
 OS J07115988-A.  
 AC 09-MAY-1995.  
 AC 26-OCT-1993; JP-267644.  
 AC (NIBS ) JAPAN TOBACCO INC.  
 AC WPI: 95-202847/27.  
 AC N-PSDB: Q95310.  
 AC Preparation of soluble membrane proteins - for their use in antibody  
 PT production for the treatment and prevention of related diseases  
 PS Claim 10; Pages 44-46; 51pp; Japanese.  
 CC R78616 is the protein prod. of the expression vector pME18S/MPD-1  
 CC extracellular region-AIC2A. The expression vector was used for  
 CC the prodn. of recombinant soluble membrane proteins. The proteins can  
 CC be used in antibody prodn. for the treatment and prevention of related  
 CC diseases.  
 CC Sequence 596 AA;  
 SQ

Query Match 7.3%; Score 115; DB 1; Length 596;  
 Best Local Similarity 32.4%; Pred. No. 1.46e-01;  
 Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 517 YIDHFFQVQYKKSES-WKDSKTENLGRVNSMDLPQLEPDTSYCARVRVKPISD-YD-GI 573  
 QY 143 YGDLLEYVQYRSPFTEW-QSKQENTCNV-TIE-G-LDAEKCYSEFWVRKAMEDVYGPDT 198  
 Db 574 W-SEWSNEYTW 583  
 QY 199 YPSDWSEVTCW 209

RESULT 11

R92526 standard; Protein; 600 AA.  
 AC R92526;  
 AC 06-SEP-1996 (first entry)  
 DE Fas antigen #1.  
 DE Fas antigen; immunoassay; monoclonal antibody; autoimmune disease; SLE;  
 KW rheumatoid arthritis; serum; systemic lupus erythematosus.  
 OS Synthetic.  
 OS Location/Qualifiers  
 FH Key 1..16 "signal peptide"  
 FT peptide 17..600  
 FT protein /note= "mature Fas antigen #1"  
 FT WO9601277-A1.  
 AC 18-JAN-1996.  
 AC 03-MAR-1995; J00349.  
 AC 06-JUL-1994; JP-154706.  
 AC 14-FEB-1995; JP-025637.  
 AC (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.  
 AC (NIBS ) JAPAN TOBACCO INC.  
 AC Hachiya T, Noguchi J, Yonehara S;  
 AC WPI: 96-087635/09.  
 AC N-PSDB: T16300.  
 AC Immunoassay method for soluble Fas antigen in body fluids - for  
 PT diagnosis of auto-immune diseases such as rheumatoid arthritis and  
 PT systemic lupus erythematosus  
 PS Claim 13; Page 73-77; 124pp; Japanese.  
 CC R92526 and R92527 represent soluble Fas antigens. These soluble Fas  
 CC antigen is included in the immunoassay kit of the invention. The kit is  
 CC for the assay of soluble Fas antigen and contains an immobilised  
 CC anti-soluble Fas monoclonal antibody, as well as one of these standard  
 CC soluble Fas antigens. The assay is simple and has high accuracy, high  
 CC sensitivity, and is capable of assaying a number of different specimens  
 CC at the same time. The immunoassay is used on biological samples (such as  
 CC serum) and is useful for diagnosis of autoimmune diseases such as  
 CC rheumatoid arthritis or systemic lupus erythematosus (SLE).  
 CC Sequence 600 AA;  
 SQ

Query Match 7.3%; Score 115; DB 1; Length 600;  
 Best Local Similarity 32.4%; Pred. No. 1.46e-01;  
 Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 521 YIDHFFQVQYKKSES-WKDSKTENLGRVNSMDLPQLEPDTSYCARVRVKPISD-YD-GI 577  
 QY 143 YGDLLEYVQYRSPFTEW-QSKQENTCNV-TIE-G-LDAEKCYSEFWVRKAMEDVYGPDT 198  
 Db 578 W-SEWSNEYTW 587  
 QY 199 YPSDWSEVTCW 209

RESULT 12  
 ID R78610 standard; Protein; 600 AA.  
 AC R78610;  
 AC 19-FEB-1996 (first entry)  
 DE Expression vector pME18S/hFas, EXT-AIC2A protein prod.  
 DE Expression vector; pME18S/hFas, EXT-AIC2A; human Fas antigen;  
 KW extracellular; region; AIC2A; soluble membrane protein;  
 KW antibody production; diseases; treatment; prevention.  
 OS Homo sapiens.  
 OS Location/Qualifiers  
 FH Key 1..16  
 FT peptide /label= sig\_peptide  
 FT peptide 17..600  
 FT peptide /label= mat\_peptide  
 FT J07115988-A.  
 AC 09-MAY-1995.  
 AC 26-OCT-1993; JP-267644.  
 AC (NIBS ) JAPAN TOBACCO INC.  
 AC WPI: 95-202847/27.  
 AC N-PSDB: Q95301.  
 AC Preparation of soluble membrane proteins - for their use in antibody  
 PT production for the treatment and prevention of related diseases  
 PT

R92527;  
06-SEP-1996 (first entry)  
Fas antigen #2  
Fas; antigen; immunoassay; monoclonal antibody; autoimmune disease; SLE;  
rheumatoid arthritis; serum; systemic lupus erythematosus.

for screening agonists and antagonists of human G  
diagnosing myeloid leukaemia

US-09-376-430-2-04.rag

Thu May 11 06:49:31 2000

PS Claim 2; Page 15-18; 26pp; English.  
 CC The nucleic acid encoding the human GM-CSF beta-chain is isolated  
 CC from a cDNA library prepared from POLY(A)+ RNA from TF-1 cells. The  
 CC high affinity human GM-CSF receptor (Kd < 1 nM) can be used for  
 CC screening candidate GM-CSF agonists and antagonists e.g. for  
 CC treating myeloid leukaemias.  
 SQ Sequence 897 AA;

Query Match 7.1%; Score 113; DB 1; Length 897;  
 Best Local Similarity 31.9%; Pred. No. 2.19e-01;  
 Matches 22; Conservative 10; Mismatches 34; Indels 3; Gaps 3;  
 DB 372 FEIQYRKDTAT-WKDSKTTLQNAHSMALPALEPSTRYWARVVRTSRTGYNGIWESE 430  
 QY :||||| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|  
 148 YEQYRSPDTEW-QSKQENTCNVTIEGLDA-EKCYFWVRVVRKAMEDVYGPDTYPSDASE 205  
 DB 431 ARSWDTESV 439  
 QY :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|  
 206 VTCWQGEI 214

Search completed: Wed May 10 11:35:56 2000  
 Job time : 13 secs.





PARENTS' APPLICATION DATA:

```
CC CC APPLICATION NUMBER: US/08/595,974
CC FILING DATE: 06-FEB-1996
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA: US 08/052,205
CC APPLICATION NUMBER: 435
CC FILING DATE: 22-APR-1993
CC APPLICATION NUMBER: JP 104947/1992
CC FILING DATE: 23-APR-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Obion, No. 570560Bman F.
CC REGISTRATION NUMBER: 24,618
CC REFERENCE/DOCKET NUMBER: 10-615-0X
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703) 413-3000
CC TELEFAX: (703) 413-2220
CC TELEX: 248855 OPAT UR
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 252 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 252 AA; 29305 MM; 347903 CN;
SQ
Query Match 12.3%; Score 195; DB 1; Length 252;
Best Local Similarity 25.9%; Pzed No. 1,13e-08;
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11.
Db 59 EVQCFEYVENVKNCWNSSEQPINLLHWYKRSDNDKYOKCSHYFSSEITSGCOLQ 118
Oy 31 QIQIYVENLETVOYNNAKYSR-INLIFHYRF-NGD-EAYDCGNTNVLQGHGTSCILD 87
Db 119 KKEHLHQTFFVOLQDDPREPRQAOTMLKLQNLVTPMAPENLTLEKLSQSOLELNWNRFF 178
Oy 88 AERDDLILFESR-NGTR-PVFPLASMM-VYL-KPSPKHVRF-SWGQDAVTVCSPOLS 142
Db 179 LNHGLEHLVQRTDMDSHTEOSVDYRHKEFLPSVDGCKRYEFYRSR-FNLGSAGHW 237
Oy 143 YGD-LLYEVQYRRPFDETW-QSKQENTCVTIGEDAKENCISFWVRKAMDEVYPDITYP 200
Db 238 SEMSHPIHW 246
Oy 201 SDMSEVTCW 209
RESULT 4 STANDARD: PRT; 252 AA.
XX ID US-08-052-205-9
AC xxxxxx
DB Sequence 9, Application US/08052205
CC Patent No. 5510259
CC GENERAL INFORMATION:
CC APPLICANT: SUGAMURA, KAZUO
CC APPLICANT: TAKASHITA, TOSHITAKU
CC APPLICANT: ASAO, HIROKOBO
CC APPLICANT: NAKAMURA, MASATAKA
CC APPLICANT: SHIMAMURA, TOSHIHO
CC APPLICANT: SUZUKI, MANABU
CC APPLICANT: HAMURO, JUNJI
CC TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE
CC NUMBER OF SEQUENCES: 21
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
CC ADDRESS: P.C.
CC STREET: 1755 S. Jefferson Davis Highway, Suite 400
CC CITY: Arlington
CC STATE: Virginia
CC COUNTRY: U.S.A.
```

	CC	ZI:	22202	
	CC	COMPUTER READABLE FORM:		
	CC	MEDIUM TYPE:	Floppy disk	
	CC	COMPUTER:	IBM PC compatible	
	CC	OPERATING SYSTEM:	PC-DOS/MS-DOS	
	CC	SOURCE:	Patentin Release #1.0, Version #1.25	
	CC	CURRENT APPLICATION NUMBER:	US/08/052.205	
	CC	APPLICATION DATE:	19930422	
	CC	CLASSIFICATION:	435	
	CC	PRIOR APPLICATION DATA:		
	CC	APPLICATION NUMBER:	JP 104947/1992	
	CC	FILING DATE:	23-APR-1992	
	CC	ATTORNEY/AGENT INFORMATION:		
	CC	NAME:	Oblon, No. 351025man F.	
	CC	REGISTRATION NUMBER:	24,618	
	CC	REFERENCE//DOCKET NUMBER:	10-615-OX	
	CC	TELECOMMUNICATION INFORMATION:		
	CC	TELEPHONE:	(703) 413-3000	
	CC	TELEX:	248855 OPAT UR	
	CC	INFORMATION FOR SEQ ID NO:	9:	
	CC	SEQUENCE CHARACTERISTICS:		
	CC	LENGTH:	252 amino acids	
	CC	TYPE:	AMINO ACID	
	CC	TOPOLOGY:	linear	
	CC	MOLECULE TYPE:	protein	
	CC	SEQUENCE	252 AA; 29505 MW; 347903 CN;	
	DB	Query Match	12.3%; Score 195; DB 1; Length 252;	
	DB	Best Local Similarity	25.9%; Pred. No. 1.13e+08;	
	DB	Matches	49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;	
Dy	59	EVOGCFENNENYMCNCTNSSEPOPNLTLHWKNSDNDKVQCSHYLFSEITSGCOLQ	118	
Dy	31	: :	87	
Dy	119	KREIHLYQTIVVLODPRRPRQRQAQMELKLQVLVPMPNPENTLHKLSOQLDELMMNRF	178	
Dy	88	AEGDDDLIYSIR-NGT-H-PVTASRW-M-VYL-KPSFKHYRF-SWHDAYTVYCGLS	142	
Dy	179	LNRCLHLVORTDMDSHTESVDYRHFFSLPSYVGOKRYTFEVRSR-FNLICSAOHM	237	
Dy	143	YGD-LIVEYEVSFPTDW-QSKOEINTCVTIEGLDAECISFWVRVKAMEDYVGPDTYP	200	
Dy	238	SEMSPPIHM 246		
Dy	201	SDMSEVTGW 209		
RESULT	5	STANDARD:	PRT; 347 AA.	
ID	XX	US-08-595-974-7		
AC	XX	xxxxxx		
DE	XX			
Sequence 7,	Application US/08595974			
Patent no.	5705608			
GENERAL INFORMATION:				
APPPLICANT:	SUGAMURA, KAZUO			
APPPLICANT:	TAKEISHITA, TOSHIAZU			
APPPLICANT:	ASAHO, HIROHARU			
APPPLICANT:	NAKAMURA, MASATAKA			
APPPLICANT:	SHIMAMURA, TOSHIO			
APPPLICANT:	SUZUKI, MANABU			
APPPLICANT:	HANBONO, JUNJI			
TITLE OF INVENTION:	HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE			
NUMBER OF SEQUENCES:	21			
CROSSREFERENCE ADDRESS:				
ADDRESSES:	OBLON, SELIVAK, MCCLELLAND, MAIER & NEUSTADT,			

US-09-376-430-2-04.rai

CC ADDRESS: P. C.  
CC STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CC CITY: Arlington  
CC STATE: Virginia  
CC COUNTRY: U.S.A.  
CC ZIP: 22202  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/595,974  
CC FILING DATE: 06-FEB-1996  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/052,205  
CC FILING DATE: 22-APR-1993  
CC APPLICATION NUMBER: JP 104947/1992  
CC FILING DATE: 23-APR-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: O'Brien, No. 5705608man F.  
CC REGISTRATION NUMBER: 24,618  
CC REFERENCE/DOCKET NUMBER: 10-615-0X  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (703) 413-3000  
CC TELEFAX: (703) 413-2220  
CC TELEX: 248855 OPAT UR  
CC INFORMATION FOR SEQ ID NO: 7:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 347 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: Protein  
CC SEQUENCE 347 AA; 35920 MW; 666553 CN;

CC APPLICANT: SHIMAMURA, TOSHIO  
CC APPLICANT: SUZUKI, MANABU  
CC APPLICANT: HANURO, JUNJI  
CC TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE  
CC NUMBER OF SEQUENCES: 21  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
CC ADDRESSEE: P.C.  
CC STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CC CITY: Arlington  
CC STATE: Virginia  
CC COUNTRY: U.S.A.  
CC ZIP: 22202  
CC  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/052,205  
CC FILING DATE: 19930422  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: JP 104947/1992  
CC FILING DATE: 23-APR-1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Obion, No. 5510259man F.  
CC REGISTRATION NUMBER: 24,618  
CC REFERENCE/DOCKET NUMBER: 10-615-0X  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (703) 413-3000  
CC TELEFAX: (703) 413-2220  
CC TELEX: 248855 OPAT UR  
CC INFORMATION FOR SEQ ID NO: 7:

[illegible]

MOLECULE TYPE: protein  
SEQUENCE . 347 AA; 39920 MW; 666553 CN;

QY 88 AEQRDDILYFSR-NGTH-PVETASRMM-VYYL-KPSSPKHYRF-SWHDQAVTTCSDLS 144

D**b**

```
      37 EVQCFVFNVEIMNCTWNSSEPOPTNLTLHYWKNSNDKQKCSHLYLSEEITSCQLQ 96  
::+ : + :: : + : + : + : + : + : + : + : + : + : + : + : + : + : +  
          Cysteine / / / Indels 11; Gaps 11
```

Db	216	SEWSHPITHW	224
		:	
QY	201	SDWSEVTCW	209

88 AEORDILYFSIR-NGTH-PVFATSRMM-VYTL-KPSSPKHVRF-SWQDAVTVCSDL 142

0S-08-032-205-/	STANDARD;	PRT;	347 AA..
XXXXXX			
XAC			
XX			

Db	216	SEWSHPITHW	224
QY	201	SDMSEVTCW	209

Sequence 7, Application US/0805220  
Patent No. 5510259

STANDARD; PRT; 369 AA

APPLICANT: ASAO, HIRONOBU  
APPLICANT: NAKAMURA, MASATAKA

Sequence 4: Application ITS/0850507

Sequence 4, Application US/08595974  
Patent No. 5705608







US-09-376-430-2-04.rai

[illegible]

RESULT	ID	US-08-609-572-2	STANDARD:	PRT:	363 AA.
XX	AC	xxxxxx			
XX	DT				
XX	DE	Sequence 2, Application us/08609572			
CC	CC	Sequence 2, Application us/08609572			
CC	CC	Patent No. 5710023			
CC	CC	GENERAL INFORMATION:			
CC	CC	APPLICANT: Collins, Mary			
CC	CC	APPLICANT: Donaldson, Debra			
CC	CC	APPLICANT: Fitz, Lori			
CC	CC	APPLICANT: Neben, Tamlyn			
CC	CC	APPLICANT: Whilters, Matthew			
CC	CC	APPLICANT: Wood, Clive			
CC	CC	TITLE OF INVENTION: CYTOKINE RECEPTOR CHAIN			
CC	CC	NUMBER OF SEQUENCES: 9			
CC	CC	CORRESPONDENCE ADDRESS:			
CC	CC	ADDRESSEE: Genetics Institute, Inc.			
CC	CC	STREET: 87 Cambridgepark Drive			
CC	CC	CITY: Cambridge			
CC	CC	STATE: MA			
CC	CC	COUNTRY: USA			
CC	CC	ZIP: 02140			
CC	CC	COMPUTER READABLE FORM:			
CC	CC	MEDIUM TYPE: Floppy disk			
CC	CC	COMPUTER: IBM PC compatible			
CC	CC	OPERATING SYSTEM: PC-DOS/MS-DOS			
CC	CC	SOFTWARE: PatentIn release #1.0, Version #1.25			
CC	CC	CURRENT APPLICATION DATA:			
CC	CC	APPLICATION NUMBER: US/08/609,572			
CC	CC	FILING DATE:			
CC	CC	CLASSIFICATION: 530			
CC	CC	ATTORNEY/AGENT INFORMATION:			
CC	CC	NAME: Brown, Scott A. 32,724			
CC	CC	REGISTRATION NUMBER: 32,724			
CC	CC	REFERENCE/DOCKET NUMBER: G15268			
CC	CC	TELECOMMUNICATION INFORMATION:			
CC	CC	TELEPHONE: (617) 498-8224			
CC	CC	TELEFAX: (617) 876-5851			
CC	CC	INFORMATION FOR SEQ ID NO: 2:			
CC	CC	SEQUENCE CHARACTERISTICS:			
CC	CC	LENGTH: 383 amino acids			
CC	CC	TYPE: amino acid			
CC	CC	TOPOLOGY: linear			
CC	CC	MOLECULE TYPE: protein			
CC	CC	SEQUENCE 383 AA; 44482 MW; 795803 CN;			
SO	SO				
DB	DB	127	8.2%; Score 130; DB 1; Length 383;		
DB	DB	127	Best Local Similarity 30.2%; Pred. No. 6,92e-03;		
DB	DB	127	Matches 26; Conservative 22; Mismatches 32; Indels 6; Gaps 5		
DB	DB	127	GGAGAGAGV-IGIITFNFETVQVWMSK--YSRTNITFHYRFG-DEAYIDCQTNVLQ 78		

DB	186	EKNVGRKLSNIDSSDYKDFICVNGS	211
OY	79	GHTSGCLDAEQRDIL-YFSIRNGT	103
RESULT	12		
ID	US-07-960-389-2	STANDARD:	PRT: 897 AA.
XX	xxxxxx		
DT			
DE	Sequence 2, Application US/07960389		
CC	Sequence 2, Application US/07960389		
CC	Patent No. 5703611		
CC	GENERAL INFORMATION:		
CC	APPLICANT: HAYASHIDA, Kasuhiko;		
CC	TITLE OF INVENTION: Human GM-CSF Receptor Component		
CC	NUMBER OF SEQUENCES: 2		
CC	CORRESPONDENCE ADDRESSES:		
CC	ADDRESSEE: Schering-Plough Corporation		
CC	STREET: 2000 Galloping Hill Road		
CC	CITY: Kenilworth		
CC	STATE: New Jersey		
CC	COUNTRY: USA		
CC	ZIP: 07033		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy Disc		
CC	COMPUTER: Apple Macintosh		
CC	OPERATING SYSTEM: System Software 7.1		
CC	SOFTWARE: Microsoft Word 5.1a		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/07/960,389		
CC	FILING DATE: 07-JAN-1993		
CC	CLASSIFICATION: 800		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: 554,745		
CC	FILING DATE: 18-JUL-1990		
CC	APPLICATION NUMBER: PCT/US 91/04846		
CC	FILING DATE: 16-JUL-1991		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Blasdale, John H. C.		
CC	REGISTRATION NUMBER: 31,895		
CC	REFERENCE/DOCKET NUMBER: DX01430		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: (908) 298-2802		
CC	TELEFAX: (908) 298-3388		
CC	INFORMATION FOR SEQ ID NO: 2:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 897 amino acids		
CC	TYPE: amino acid		
CC	STRANDEDNESS: single		
CC	TOPOLOGY: linear		
CC	MOLECULE TYPE: protein		
CC	FEATURES:		
CC	OTHER INFORMATION: Human GM-CSF receptor; Signal Sequence: -17 to -1		
SO	SEQUENCE 897 AA: 97323 MW: 4443497 CN:		
DQ	Query Match 7.1%; Score 113; DB 1; Length 897;		
DQ	Best Local Similarity 31.9%; Pred. No. 1,85e-01;		
OY	Matches 22; Conservative 10; Mismatches 34; Indels 3; Gaps 3;		
DB	372 FEILOYRDKDT-KWDSKTFLONAHSMALDELPESTRYMARVRYSRTGYINGIWMSWE	430	
OY	148 YEVQYRSRPFTLW-QSKQENTCNVTIEGLDA-EKCYSEFWVRKAMEDVDYGPDTPSDMSE	205	
DB	431 ARSWDTESV	439	
OY	206 VTQMGRGEI	214	
RESULT	13		

ID US-08-673-789-4 STANDARD; PRI; 982 AA.  
AC xxxxxx

Sequence 4, Application US/08673789

Sequence 4, Application US/08673789  
Patent No. 5814479

GENERAL INFORMATION:

APPLICANT: ZHOU, RENGING; SCHULZ, NICHOLAS,  
APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,  
APPLICANT: GEORGE, F.

TITLE OF INVENTION: BSK RECEPTOR LIKE

TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR  
TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESSES:

ADDRESS: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

OPERATING SYSTEM: IBM PC COMPATIBLE

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/673,789

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/177,812

FILING DATE: 04-JAN-1994

ATTORNEY/AGENT INFORMATION:

NAME: CAROL M. GRUPEL

REGISTRATION/DOCKET NUMBER: 37,341

REFERENCE/DOCKET NUMBER: 2026-4105

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 982

TYPE: AMINO ACID

STRANDEDNESS: UNKNOWN

TOPOLOGY: UNKNOWN

SEQUENCE 982 AA; 109827 MW; 5153844 CN;

Query Match

Best Local Similarity 6.2%; Score 99; DB 2; Length 982;  
Matches 18; Conservative 15; Mismatches 26; Indels 3; Gaps 3;

Db 459 PEHPGIIIDVEVYKQOETSTYILRAKSTNVTISGLKPDITYEOTRART-PAARYET 517  
OY 139 SDSLSTGDL-VEVOY-RSPFDTEWQSKOENTCNVTEIGLDAKCYSEFWVRAMEDYGP 196  
Db 518 SS 519  
OY 197 DT 198

DE Sequence 4, Application US/08609572

Sequence 4, Application US/08609572  
Patent No. 5710023

GENERAL INFORMATION:

APPLICANT: COLLINS, Mary

APPLICANT: Donaldson, Debra

APPLICANT: Pitz, Lori

APPLICANT: Neben, Tamlyn

APPLICANT: Whitters, Matthew

APPLICANT: Wood, Clive

TITLE OF INVENTION: CYTOKINE RECEPTOR CHAIN

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:

ADDRESS: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: USA  
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/609,572

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Brown, Scott A.

REGISTRATION/DOCKET NUMBER: 32,724

REFERENCE/DOCKET NUMBER: G15268

TELEPHONE: (617) 498-8224

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 380 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: Protein

SEQUENCE 380 AA; 44176 MW; 816181 CN;

Query Match

Best Local Similarity 5.9%; Score 93; DB 1; Length 380;  
Matches 19; Conservative 19; Mismatches 23; Indels 5; Gaps 4;

Db 133 OGIPETKQDMQCVYVNMQYLCSMRGIVGLDNTNXLNFYIEGLDHAL-CCVDYIKAD 191  
OY 23 OGGAAGVQ-IGIIFYNLETVYTNASK--YSRNLFHYRNG-DEAYDOCTNVLLOE 78  
Db 192 GONIGC 197  
OY 79 GHTSGC 84

RESULT 15

ID US-08-485-216-2 STANDARD; PRI; 446 AA.  
AC xxxxxx

Sequence 2, Application US/08485216

Sequence 2, Application US/08485216  
Patent No. 5695960

GENERAL INFORMATION:

APPLICANT: Chan, Voon Loong  
APPLICANT: Hani, Eric K  
TITLE OF INVENTION: HIPURICASE GENE

Thu May 11 06:49:32 2000

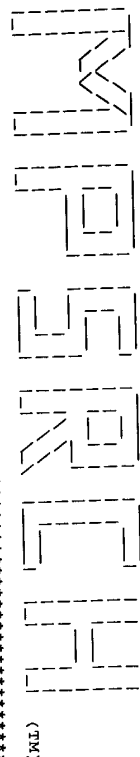
US-09-376-430-2-04.rtf

Page 9

CC NUMBER OF SEQUENCES: 2  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Merchant & Gould  
CC STREET: 3100 No. 5695960west Center, 90 S. 7th Street  
CC CITY: Minneapolis  
CC STATE: Minnesota  
CC COUNTRY: U.S.A.  
CC ZIP: 55402  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/485,216  
CC FILING DATE: 07-JUN-1995  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO 94/26907  
CC FILING DATE: 13-MAY-1994  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Mueller, Douglas P.  
CC REGISTRATION NUMBER: 30,300  
CC REFERENCE/DOCKET NUMBER: 7933.290S02  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 612/332-5300  
CC TELEFAX: 612/332-9081  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 446 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: Protein  
CC SEQUENCE 446 AA: 50277 MW: 986088 CN;  
SQ  
Query Match 5.6%; Score 89; DB 1; Length 446;  
Best Local Similarity 31.3%; Pred. No. 1.51e+01;  
Matches 10; Conservative 7; Mismatches 15; Indels 0; Gaps 0;  
DB 6 RIPPLYHLLMGSKSWSDKTRHFTLGFYIF 37  
QY 31 QIQLIFNLETVQVTWNAKSKYKRTNLFHIRE 62

Search completed: Wed May 10 11:37:09 2000  
Job time : 57 secs.





Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MSPch.p protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:35:06 2000; Maspar time 15.40 Seconds  
640.208 Million cell updates/sec

Tabular output not generated.

Title: >US-09-376-430-2  
Description: (23-231) from US09376430A.pep (4 of 25)  
Perfect Score: 1586  
Sequence: 1 OGGAAGVQIQITTFNLETV.....GEIRACAEPTPPPKLSK 209

Scoring table:  
PAM 150  
Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: p1r62  
1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 43.035; Variance 76.141; scale 0.565

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	210	13.2	373	A55718	interleukin-2 recepto	1.68e-21
2	195	12.3	369	A42565	interleukin-2 recepto	1.46e-18
3	188	11.9	369	A42565	interleukin-2 recepto	3.28e-17
4	115	7.3	878	A40081	interleukin-3 recepto	3.45e-04
5	113	7.1	897	A39255	cytokine receptor com	7.16e-04
6	109	6.9	435	E71350	probable aspartate am	3.02e-03
7	108	6.8	896	A37820	cytokine receptor com	6.91e-02
8	100	6.3	638	S04530	somatotropin receptor	5.98e-02
9	100	6.3	638	A33991	somatotropin receptor	5.98e-02
10	99	6.2	4436	E71086	somatotrophic protein	1.89e-01
11	99	6.2	6839	S57242	hypothetical protein	1.89e-01
12	97	6.1	302	S50579	hypothetical protein	2.63e-01
13	97	6.1	311	S77803	hypothetical protein	2.63e-01
14	96	6.1	638	S12136	hypothetical protein	5.05e-01
15	93	5.9	173	T15695	hypothetical protein	5.05e-01
16	93	5.9	173	T15695	hypothetical protein	5.05e-01
17	94	5.9	638	B28176	somatotropin receptor	5.05e-01
18	94	5.9	26926	1	aldose 1-epimerase-11	9.62e-01
19	92	5.8	358	T07719	glutamate-aminonua-11	9.62e-01
20	92	5.8	361	JN0716	novel antigen recepto	1.84e+00
21	92	5.8	684	S60266	hypothetical protein	1.84e+00
22	90	5.7	376	D72493	ribosomal protein S6	1.84e+00
23	90	5.7	629	A30001		

24	89	5.6	255	2	E72226	conserved hypothetical	2.48e+00
25	89	5.6	1285	2	A42150	P-glycoprotein pgp1 -	2.48e+00
26	88	5.5	137	2	J02325	glutathione S-transferase	3.38e+00
27	87	5.5	416	1	BVEBHA	glutathione S-transferase	4.60e+00
28	87	5.5	430	2	UC2334	lanosterol 14alpha-de	4.60e+00
29	87	5.5	486	2	UC4758	lanosterol 14alpha-de	4.60e+00
30	87	5.5	503	2	UC4758	lanosterol 14alpha-de	4.60e+00
31	88	5.5	562	1	Q08B55	phosphotransferase (E	4.60e+00
32	87	5.5	634	1	S33359	somatotropin receptor	4.60e+00
33	87	5.5	873	1	Q08B55	somatotropin receptor	4.60e+00
34	87	5.5	918	1	A36357	membrane glycoprotein	4.60e+00
35	88	5.5	935	1	I41193	outer membrane protein	4.60e+00
36	87	5.5	939	2	I41197	outer membrane protein	4.60e+00
37	88	5.5	983	2	B45583	receptor tyrosinase kin	4.60e+00
38	87	5.5	1354	2	T13930	myb-like protein	5.24e+00
39	86	5.4	43	2	T06456	probable acylaminoacy	8.45e+00
40	86	5.4	426	2	B71408	intermediate filament	6.24e+00
41	85	5.4	458	2	S09228	2-oxoacid-ferredoxin	6.24e+00
42	86	5.4	632	1	JC4919	exodeoxyribonuclease	8.45e+00
43	86	5.4	732	2	A72118	Syl1 protein - yeast	8.45e+00
44	85	5.4	902	2	S49931	proteoglycan - fruit fly	
45	85	5.4	6658	2	T13931		

## ALIGNMENTS

RESULT 1  
ENTRY 1  
TITLE interleukin-2 receptor gamma chain precursor - dog  
ORGANISM Canis lupus familiaris #common\_name dog  
DATE 18-Aug-1995 #sequence\_rev18-Aug-1995 #text\_change 23-Jul-1999

ACCESSIONS  
REFERENCE  
#authors  
#journal  
#title  
#cross-references M01D:95130114  
#accession A55718  
#status preliminary  
#molecule\_type mRNA  
#residues 1-373 #label HEN  
#cross-references GB:U04361; NID:9517411; PIDN:AAC48403.1; PID:9517412  
#superfamily interleukin-2 receptor gamma chain  
KEYWORDS  
#length 373 #molecular\_weight 42516 #checksum 7960  
SUMMARY

Query Match 13.2%; Score 210; DB 2; Length 373;  
Best Local Similarity 25.9%; Pred. No. 1.68e-21;  
Matches 49; Conservative 56; Mismatches 73; Indels 11; Gaps 10;

Db	59	EVGCEPVENYMNCTMNSSEPRPTLTHVYKNSNDKVOECGYLFSREVTAGCWLQ	118
Qy	31	QDITTFNLETVQYVWNAKYSR-TLTFHYRF-NGDEA-YDQCTNYLLDQCHTSGCLLD	87
Db	119	KEIHLVETFFVQVLDPRPRRQSTOKLQNIYVWAEENITLHNSQSLTSMNSRH	178
Qy	88	AEQPD-DILYSIKNGHPVETAS-RMNVYTL-KRSPKHYRF-SWHDATVVCSDLS	142
Db	179	LDRCLEHVQVRSQDWSVTEQSDVDRHSFUSVGOFFYFRVRSR-NPLCGSAQW	237
Qy	143	YGD-LTVEVQVRSFDEW-QSKQENICNTLIGDAEKYSFWVAKMEDYVGPDTYP	200
Db	238	SEWSHPITHW 246	
Qy	201	SDWSEVTCW 209	
RESULT 2		A42565	#type complete
ENTRY			

```

TITLE      Interleukin-2 receptor gamma chain - human
ORGANISM   #formal_name Homo sapiens #common_name man
DATE       04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change
ACCESSIONS A42565; A46591; I54332
REFERENCE   A42565
#authors   Takeshita, T.; Asao, H.; Ohtani, K.; Ishii, N.; Kumaki, S.;
#journal   Tanaka, N.; Munakata, H.; Nakamura, M.; Sugamura, K.
#title     Cloning of the gamma chain of the human IL-2 receptor.
#cross-references MIM:92335883
#accession A42565
#status    Preliminary; not compared with conceptual translation
#molecule_type nucleic acid; protein
#residues  1-369 #label TAK
#cross-references GB:D11086; NID:9303611; PIDN:BAA01857.1;
#experimental_source MOLT beta lymphoid cells
#note      sequence extracted from NCBI backbone (NCBI:109167)
REFERENCE   A46591
#authors   Noguichi, M.; Adelstein, S.; Cao, X.; Leonard, W.J.
#journal   J Biol. Chem. (1993) 268:13601-13608
#title     Characterization of the human interleukin-2 receptor gamma
#cross-references MIM:93293887
#accession A46591
#status    Preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues  1-369 #label RES
#cross-references GB:J12183; NID:9307056; PIDN:AAA59145.1; PID:9307058
REFERENCE   I54332
#authors   Puck, J.M.; Deschenes, S.M.; Porter, J.C.; Dutra, A.S.;
#journal   Brown, C.J.; Willard, H.F.; Henthorn, P.S.
#title     Hum. Mol. Genet. (1993) 2:1099-1104
#cross-references GB:J12183; NID:9307056; PIDN:AAA59145.1; PID:9307058
#note      The interleukin-2 receptor gamma chain maps to Xq13.1 and is
#note      mutated in X-linked severe combined immunodeficiency,
#note      SCIDX1.
#cross-references MIM:94004847
#accession I54332
#status    Preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues  1-369 #label RE2
#cross-references GB:J12546; NID:9349631; PIDN:AAC37524.1; PID:9349632
GENETICS
#gene      GDB:112RG; SCIDX1; IMD4
#map_position Xq13.1; Xq13.1
#introns   39/1; 90/2; 152/1; 198/3; 253/1; 285/2; 308/3
#note      defects are associated with an X-linked form of severe
#note      combined immunodeficiency
#classification #superfamily interleukin-2 receptor gamma chain
#keywords       cytokine receptor; duplication; immunodeficiency; severe
#keywords       combined immunodeficiency; transmembrane protein
#length      369 #molecular_weight 42287 #checksum 7850
SUMMARY
Query Match 12.3%; Score 195; DB 2; Length 369;
Best Local Similarity 25.9%; Pred. No. 146e-18;
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;
Db 59 EVOCFENYEMNCTNSSEPOPTVLTLYHWKNSDNDKVKCSHYLSEETSCQO 118
OY 31 QOITFNEIYVQVWNAKYSR-TWLTFRYF-NGD-EAYDCTVYLQEGHTSCLD 87
Db 119 KEIHLVQTFVQLODPRPRROATOMLKLQNVIPWAPENLTLHLSQQLNNNFR 178
OY 88 AEORDIYFSIR-NGTH-PVFTASRM-VYTL-KPSSKHYRF-SWMDAVYVTSIDS 142
Db 179 LNHCEHYVORTMDHSTEOSVDYRHRFSPVSGOKRYFRYSR-FNPLCSAQHM 237
OY 143 YGD-LLYEVQSRSPDTEM-OSKOENTCNVTEGLDAEKQYSFWVRVAMEDVYGPDIY 200
Db 238 SEWHPPIHM 246
I:|:|

```

---

```

OY 201 SDMSVTCW 209
RESULT      3
ENTRY       149280 #type complete
TITLE       Interleukin-2 receptor gamma chain precursor - mouse
ORGANISM    #formal_name Mus musculus #common_name house mouse
DATE        27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change
ACCESSIONS I49280; A47514; JN0592; JN0775; S37582; I53398
REFERENCE   I49280
#authors   Cao, X.; Kozak, C.A.; Liu, Y.
#journal   Proc. Natl. Acad. Sci. U.S.A. (1993) 90:8464-8468
#title     Characterization of cDNAs encoding the murine interleukin 2
#cross-references MIM:93391374
#accession I49280
#status    Preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues  1-369 #label CAO
#cross-references EMBL:U21795; NID:9727349; PIDN:AAA64279.1;
#accession A47514
#status    Translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues  1-369 #label RE2
#cross-references GB:U20048; NID:9404067; PIDN:AAA39286.1; PID:9404068
REFERENCE   JN0592
#authors   Kumaki, S.; Kondo, M.; Takeshita, T.; Asao, H.; Nakamura, M.;
#journal   Sugamura, K.
#title     Biochem. Biophys. Res. Commun. (1993) 193:356-363
#cross-references GB:U20048; NID:9404067; PIDN:AAA39286.1; PID:9404068
#note      Cloning of the mouse interleukin 2 receptor gamma chain:
#note      Demonstration of functional differences between the mouse
#note      and human receptors.
#cross-references MIM:93277575
#accession JN0592
#status    nucleic acid sequence not shown
#molecule_type mRNA
#residues  1-369 #label KUM
#cross-references DBJ:D13365; NID:9303684; PIDN:BAA02760.1;
#cross-references GB:U20048; NID:9404067; PID:9404068
REFERENCE   JN0775
#authors   Kobayashi, N.; Nakagawa, S.; Minami, Y.; Taniguchi, T.; Kono,
#journal   T.
#title     Gene (1993) 130:303-304
#cross-references GB:U20048; NID:9404067; PIDN:AAA39286.1; PID:9404068
#accession JN0775
#status    Cloning and sequencing of the cDNA encoding a mouse IL-2
#molecule_type mRNA
#residues  1-369 #label KOB
#cross-references GB:U20048; NID:9404067; PID:9404068
#accession S37582
#status    Preliminary
#molecule_type mRNA
#residues  1-350; 'S', 352-366; 'S', 368-369 #label CHI
#cross-references EMBL:X75337
REFERENCE   I53398
#authors   Disanto, J.P.; Certain, S.; Wilson, A.; MacDonald, H.R.;
#journal   Avner, P.; Fischer, A.; de Saint Basile, G.
#title     Eur. J. Immunol. (1994) 24:3014-3018
#cross-references GB:U20048; NID:9404067; PIDN:AAA39286.1; PID:9404068
#note      The murine interleukin-2 receptor gamma chain gene:
#note      organization, chromosomal localization and expression in
#note      the adult thymus.
#cross-references MIM:95104285
#accession I53398

```



```

#status Preliminary; translated from GB/EMBL/DBD
##molecule-type DNA
##residues 1-369 ##label RES
##cross-references GB:S75852; NID:9861554; PIDN:AAH32904.1; PID:9861555
GENETICS
#gene IL-2Rgamma
#introns 39/1; 90/2; 152/1; 199/3; 254/1; 286/2; 308/3
#exons The high affinity receptor is a heterotrimer of alpha (see
#complex PIR:0HMS2), beta (see PIR:A5052) and gamma chains;
#heterodimers of alpha or beta and gamma chains are
#intermediate affinity receptors.
FUNCTION
#description receptor for interleukin-2
#pathway interleukin-2 stimulated growth and differentiation of T
#cells, B cells, NK cells, LAK cells, monocytes,
#macrophages, and oligodendrocytes
#superfamily interleukin-2 receptor gamma chain
#cytoreceptor duplication; glycoprotein; transmembrane
#protein
CLASSIFICATION
#keywords
FEATURE
#domain signal sequence #status predicted #label SIG\
#product interleukin-2 receptor gamma chain #status
#predicted #label MAT\
#domain transmembrane #status predicted #label TM\
23-369
256-284
71-75,84,96,159,
164,306
#binding site carbohydrate (Asn) (covalent) #status
#predicted
#length 369 #molecular-weight 42241 #checksum 6734
SUMMARY
#length 369 #score 11.94; DB 2; Length 369;
#similarity 26.84; Pred. No. 3,286-17;
#local similarity 52; Mismatches 75; Indels 12; Gaps 11;
Matches 51; Conservative
Query Match
Best Local Similarity 11.94; DB 2; Length 369;
Matches 51; Conservative 52; Mismatches 75; Indels 12; Gaps 11;
Db 59 EVQCEVFNTIYMCCTNSSEBPQATNLTTHRYKVSNNTFQCSHYLSKEITSCQIQ 118
QY 31 QIITFNLTLVQYTNASKYSR-TINTFHYRN-GDE-AYDQCTNVLLEQHTSCGLD 87
Db 119 KEDIQLYQTFVLODPQKQKQRAVQKLTNLONLIPAPENLTLSNLESGDLPRKSRH 178
QY 88 AEQDDILYFSIR-NGTH-EVFTLSRKM-VYLL-KPSSPHVFRS-WHODAVTV--TCSD 140
Db 179 IKERCKLQYLVQYRSNDRSMTLIVNEHPRFSLPSVDELKRYFRVRSRY-NPICSQQ 237
QY 141 LSYGDLLEYQVYRSPDLEW-QSKQENICNVITIGDLAKCYISWAVKMAEDVYGGDTY 199
Db 238 WSKWSQPYHW 247
QY 200 PSDMSEVTCW 209
ENTRY
#status 4
#title A40091 #type complete
#organism interleukin-3 receptor beta chain precursor - mouse
#accession J000001 #common_name house mouse
#residues 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change
22-Jun-1999
#accessions A40091; A43022
#references A40091
#authors Itchh. N.; Yonehara, S.; Schreurs, J.; Gorman, D.M.; Maruyama,
#journal K.; Ishii, A.; Yahara, I.; Areai, K.; Miyajima, A.
#title Cloning of an interleukin-3 receptor gene: a member of a
#distinct receptor gene family.
#cross-references MIMD:90117145
#accession A40091
#status nucleic acid sequence not shown
#molecule-type mRNA
#residues 1-878 ##label ITO
REFERENCE
#authors Gorman, D.M.
#submission Submitted to GenBank, November 1989
#accession A43022
##molecule-type mRNA

```

```

##residues 1-815,'Q',817-878 #label GOR
##cross-references GB:M29685; NID:q198342; PIDN:AAA39295.1; PID:q309406
COMMENT In mice there are two classes of high-affinity IL-3 receptors. One
beta chain also shared by high-affinity IL-5 and GM-CSF
receptors
CLASSIFICATION #superfamily interleukin-3 receptor beta chain; cytokine
receptor homology
KEYWORDS cytokine receptor; duplication; transmembrane protein
FEATURE
1-22 #domain signal sequence #status predicted #label SIG
23-878 #product.interleukin-3 receptor beta chain #status
predicted #label MAT\
23-440 #domain extracellular #status predicted #label EXT\
39-236 #domain cytokine receptor homology #label CRS1\
234-433 #domain cytokine receptor homology #label CRS2\
441-462 #domain transmembrane #status predicted #label TM\
463-878 #domain intracellular #status predicted #label INT\
#length 878 #molecular-weight 97222 #checksum 2346
SUMMARY
Query Match 7.3%; Score 115; DB 1; Length 878;
Best Local 23; Similarity 32.4%; Pred. No. 3,45e-04; Indels 8; Gaps 8;
Matches 23; Conservative 17; Mismatches 23;
Db 370 YIDHFEVOVQYKRKSES-WKSKTEMLGKVNMDIPQLEPDTGYARVYPIGD-YD-GI 426
143 YGDLLEYGVYRSPDEPTQW-SQKQENTCNV-TIE-G-IDAEKCSFWYKRAMEDVYQPD 198
OY 427 W-SEMSENYTM 436
199 YPSDMSWVTCW 209
RESULT 5
ENTRY A39255 #type complete
TITLE Cytokine receptor common beta chain precursor - human
F1000 1
F1000 1 #formal name Homo sapiens #common_name man
DATE 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change
02-Sep-1997
ACCESSION A39255
REFERENCE A39255
AUTHORS Hayashi, K.; Kitamura, T.; Gorman, D.M.; Arai, K.; Yokota
T.; Miyajima, A.
JOURNAL Proc. Natl. Acad. Sci. U.S.A. (1990) 87:9655-9659
TITLE Molecular cloning of a second subunit of the receptor for
human granulocyte-macrophage colony-stimulating factor
(GM-CSF): reconstruction of a high-affinity GM-CSF
receptor.
cross-references MOTID:91088571
Accession A39255
#molecule_type mRNA
#residues 1-897 #label HAY
cross-references GB:M3275
COMMENT The human high-affinity IL-3, IL-5, and GM-CSF receptors have
ligand-specific alpha chains and share this common beta chain.
GENETICS
#gene GDB:CSF2RB
#cross-references GDB:126838; OMTN:138981
#map position 22q13.1-22q13.1
CLASSIFICATION #superfamily interleukin-3 receptor beta chain; cytokine
receptor homology
KEYWORDS alternative splicing; cytokine receptor; duplication;
transmembrane protein
FEATURE
1-16 #domain signal sequence #status predicted #label SIG\
17-897 #product cytokine receptor common beta chain #status
predicted #label MAT\
17-443 #domain extracellular #status predicted #label EXT\
35-232 #domain cytokine receptor homology #label CRS1\
250-431 #domain cytokine receptor homology #label CRS2\
444-460 #domain transmembrane #status predicted #label TM\
461-897 #domain intracellular #status predicted #label INT\
#length 897 #molecular-weight 97323 #checksum 753
SUMMARY

```

Query Match 7.1%; Score 113; DB 1; Length 897;  
Best Local Similarity 31.9%; Pred. No. 7.16e-04;  
Matches 22; Conservative 10; Mismatches 34; Indels 3; Gaps 3;

Db 372 FEIYORRDAAT-WKDKRTETLONAHSMALPALEPSTRYARVAVRTSRGNCINSEMS 430  
148 YEVQVSPDTEW-QSKQENTCNVTIEGLDA-EKCYSFVRYKAMEDVYGPDPYPSDMS 205  
Db 431 ANSWDETV 439  
QY 206 VTCWQGEI 214

RESULT 6  
ENTRY E71350 #type complete  
TITLE Probable aspartate aminotransferase (tpart) - syphilis  
ORGANISM syphilis  
DATE 24-Jul-1998 #sequence-revision 24-Jul-1998 #text-change  
ACCESSION E71350  
REFERENCE A71250  
#authors Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin, M.; Hickey, E.K.; Clayton, R.; Ketchum, K.A.; Sodergren, E.; Hardham, J.M.; McLeod, M.P.; Salzberg, S.; Peterson, J.; Khailak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utleback, T.; McDonald, L.; Artach, P.; Bowman, C.; Cotton, M.D.; Fujii, C.; Garland, S.; Hatch, B.; Horst, H.O.; Roberts, K.; Watthey, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
#journal Science (1998) 281:375-388  
#title Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
#cross-references MWID:98332770  
#accession E71350  
#status preliminary; nucleic acid sequence not shown; translation not shown

##molecule-type DNA  
##residues 1-435 #label COL  
##cross-references GB:AE001204; GB:AE00520; NID:93322492; PID:93322493  
##experimental\_source strain Nichols

GENETICS  
#gene TP0223  
SUMMARY #length 435 #molecular-weight 47304 #checksum 4156

Query Match 6.9%; Score 109; DB 2; Length 435;  
Best Local Similarity 35.1%; Pred. No. 3.02e-03;  
Matches 20; Conservative 11; Mismatches 21; Indels 5; Gaps 5;

Db 206 AVLYICDD-ATSGEYEVASLRGFFARFOAHKNICALKIDGLTKEE-YA-WGLRV 259  
QY 133 AVYTCSDLSIGDLYEVQY-RSPFDEWQSKQENTCNVTIEGLDAEKCYSEFW-VRV 187

RESULT 7  
ENTRY A35782 #type complete  
TITLE cytokine receptor common beta chain precursor - mouse  
ORGANISM Mus musculus #common\_name house mouse  
DATE 30-Jun-1993 #sequence-revision 30-Jun-1993 #text-change  
ACCESSION A35782  
REFERENCE A35782  
#authors Gorman, D.M.; Itoh, N.; Kitamura, T.; Schreurs, J.; Yonehara, S.; Yahara, I.; Arai, K.; Miyajima, A.  
#journal Proc Natl. Acad. Sci. U.S.A. (1990) 87:5459-5463  
#title Cloning and expression of a gene encoding an interleukin 3 receptor-like protein: identification of another member of the cytokine receptor gene family.  
#cross-references MWID:90319131  
#accession A35782  
#molecule\_type mRNA

##residues 1-896 #label GOR  
#cross-references GB:W34397; NID:9191821; PIDN:AAA37204.1; PID:9309101  
COMMENT Mouse high-affinity IL-5, GM-CSF, and one class of high-affinity IL-3 receptors have ligand-specific alpha chains and share this common beta chain.  
CLASSIFICATION #superfamily Interleukin-3 receptor beta chain; cytokine receptor homology  
KEYWORDS cytokine receptor; duplication; transmembrane protein  
FEATURE 1-22  
23-896  
23-441 #domain signal sequence #status predicted #label SIG  
39-235 #product cytokine receptor common beta chain #status predicted #label MAT  
253-434 #domain extracellular #status predicted #label EXT  
442-463 #domain cytokine receptor homology #label CRS1  
464-896 #domain transmembrane #status predicted #label TMM  
SUMMARY #length 896 #molecular-weight 99110 #checksum 6643

Query Match 6.8%; Score 108; DB 1; Length 896;  
Best Local Similarity 30.6%; Pred. No. 4.32e-03;  
Matches 22; Conservative 18; Mismatches 24; Indels 8; Gaps 8;

Db 370 SEFHTROYVKKKSDS-WEDSKTENIDRAHSMDSQLEPDTSYCARVRYKPSIN-YD-G 426  
QY 142 SYGDLYEVQVRSFDEWQ-SKOEN-TCNVTE-G-LDAEKCYSEFWRYKAMEDVYGP 197  
Db 427 IW-SKWESEYTW 437  
QY 198 TYPDSWSEVTCW 209

RESULT 8  
ENTRY S04530 #type complete  
TITLE somatotropin receptor, hepatic precursor - human  
ALTERNATE\_NAMES growth hormone receptor  
CONTAINS somatotropin-binding protein, serum  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 07-Sep-1990 #sequence-revision 07-Sep-1990 #text-change  
ACCESSION S04530  
REFERENCE S04530  
#authors Leung, D.W.; Spencer, S.A.; Cachianes, G.; Hammonds, R.G.; Collins, C.; Henzel, W.J.; Barnard, R.; Waters, M.J.; Wood, W.I.  
#journal Nature (1987) 330:537-543  
#title Growth hormone receptor and serum binding protein: purification, cloning and expression.  
#cross-references MWID:88065896  
#accession S04530  
##molecule-type mRNA  
##residues 1-638 #label LEU  
##cross-references EMBL:X06562; NID:931737; PID:931738  
CLASSIFICATION #superfamily fibronectin type III repeat homology  
KEYWORDS liver; transmembrane protein  
FEATURE 1-18  
19-638  
SUMMARY #length 638 #molecular-weight 71499 #checksum 4388

Query Match 6.3%; Score 100; DB 2; Length 638;  
Best Local Similarity 32.7%; Pred. No. 6.91e-02;  
Matches 17; Conservative 11; Mismatches 22; Indels 2; Gaps 2;

Db 190 LEVLDQKVENETKRMKMDPILITTSVPYSLAKDYEVKRVNSKQNSNYG 241  
QY 146 LYEVQVRSFDEWQSKQEN-TCNVTEGLDAEKCYSEFWRYKAMEDV-YG 195

RESULT 9  
ENTRY A33991 #type complete





Thu May 11 06:49:32 2000

US-09-376-430-2-04.PR

Page 7

```
ALTERNATE_NAMES growth hormone receptor
ORGANISM #formal_name Sus scrofa domestica #common_name domestic pig
DATE 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
10-Sep-1997

ACCESSIONS
REFERENCE S12136
#authors Clift, J.A.; Wang, X.; Kopchick, J.J.
#journal Nucleic Acids Res. (1990) 18:6451
#title Porcine growth hormone receptor cDNA sequence.
#cross-references M01D:91057155
#accession S12136
#status preliminary; translation not shown
#molecule_type mRNA
#residues 1-638 #label C10
#cross-references EMBL:X54429; NID:g2037; PID:g2038
KEYWORDS transmembrane protein
SUMMARY #length 638 #molecular_weight 71144 #checksum 2806

Query Match 6.1%; Score 96; DB 2; Length 638;
Best Local Similarity 28.8%; Pred. No. 2.63e-01;
Matches 15; Conservative 15; Mismatches 20; Indels 2; Gaps 2;

DB 130 LEVELOKRVNTOKMDPVLTSTVPSYSLRLDKYEVRRVRSRQNSKRG 241
QY 146 LIXEVQYRSPFTLWQSKQEN-TCNVTLEGIDAEKCYSFVWVKAME-DVYG 195

RESULT 15
ENTRY 156563 #type complete
TITLE Interleukin-3 receptor beta-subunit - rat
ORGANISM #formal_name Rattus sp. #common_name rat
DATE 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
23-Jul-1999

ACCESSIONS
REFERENCE 156363
#authors Appel, K.; Buttini, M.; Sauter, A.; Gebicke-Haerter, P.J.
#journal J. Neurosci. (1995) 15:5800-5809
#title Cloning of rat interleukin-3 receptor beta-subunit from
cultured microglia and its mRNA expression in vivo.
#cross-references M01D:95370942
#accession 156563
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-896 #label RES
#cross-references GB:S79263; NID:g1086954; PID:AAB35068.1;
PID:g1086955

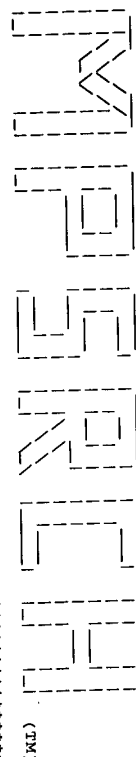
GENETICS
#gene rIL-3Rbeta
CLASSIFICATION #superfamily interleukin-3 receptor beta chain; cytokine
#receptor homology
#cytokine receptor
KEYWORDS
FEATURE 39-235
253-433 #domain cytokine receptor homology #label CRS1\
SUMMARY #domain cytokine receptor homology #label CRS2
#length 896 #molecular_weight 99504 #checksum 5352

Query Match 6.1%; Score 96; DB 2; Length 896;
Best Local Similarity 30.3%; Pred. No. 2.63e-01;
Matches 20; Conservative 17; Mismatches 21; Indels 8; Gaps 8;

DB 375 FOVOYKKKLLDR-WEDSKTENVLNHNSMDLPOLPESTGCARVRYTPE-Y-KGLW-SEW 430
QY 148 YEVQYRSPFTLWQ-SKQENTCNV-TIE-G-LDAEKCYSFVWVKAME-DVYGPPTYPEDW 203
DB 431 SNECTW.436
QY 204 SEVTCW 209

Search completed: Wed May 10 11:35:26 2000
Job time : 20 secs.
```





Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Msearch\_p protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:28:29 2000; Maspar time 100.42 Seconds  
Tabular output not generated.

Title: >US-09-376-430-2  
(23-231) from US09376430A.pep (4 of 25)

Description: 1586  
Perfect Score: 1  
Sequence: 1 QGGAGGAGVQIQIITFNLETV.....GEIRDACAEPTPPKPKLSK 209

Scoring table: PAM 150  
Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot38  
1:swissprot

Statistics: Mean 43.819; Variance 69.976; scale 0.626

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	210	13.2	373	1	CYRG CANFA	5.01e-24
2	195	12.5	369	1	CYRG HUMAN	8.64e-21
3	188	11.9	369	1	CYRG MOUSE	2.66e-19
4	187	11.8	378	1	CYRG BOVIN	4.33e-19
5	115	7.3	378	1	IL13R MOUSE	5.23e-05
6	113	7.3	378	1	CYRG HUMAN	1.16e-04
7	108	6.8	366	1	CYRG MOUSE	8.25e-04
8	100	6.3	366	1	CYRG HUMAN	1.70e-02
9	97	6.1	302	1	TEO6_TENAST	5.07e-02
10	96	6.1	302	1	GHR_PTIG	7.26e-02
11	93	5.9	380	1	IL13R_HUMAN	2.10e-01
12	94	5.9	380	1	INTERLEUKIN-13 RECEPTOR	1.48e-01
13	92	5.8	361	1	GHR_RABIT	2.98e-01
14	90	5.7	427	1	IL13R_HUMAN	5.93e-01
15	90	5.7	427	1	IL13R_HUMAN	5.93e-01
16	89	5.6	629	1	IL13R_HUMAN	8.32e-01
17	89	5.6	629	1	IL13R_HUMAN	8.32e-01
18	89	5.6	629	1	IL13R_HUMAN	8.32e-01
19	87	5.5	562	1	HEM1_SALTY	1.63e+00
20	87	5.5	562	1	HEM1_SALTY	1.63e+00
21	87	5.5	562	1	HEM1_SALTY	1.63e+00
22	87	5.5	562	1	HEM1_SALTY	1.63e+00
23	87	5.5	562	1	HEM1_SALTY	1.63e+00

Result ID	Score	Match	Length	ID	Description	Pred. No.
24	87	5.5	918	1	IL6B_HUMAN	1.63e+00
25	86	5.2	934	1	EAE2_ECOLI	1.17e+00
26	87	5.2	934	1	EAE2_ECOLI	1.17e+00
27	86	5.2	934	1	EAE2_ECOLI	1.17e+00
28	86	5.2	934	1	EAE2_ECOLI	1.17e+00
29	86	5.2	934	1	EAE2_ECOLI	1.17e+00
30	86	5.2	934	1	EAE2_ECOLI	1.17e+00
31	86	5.2	934	1	EAE2_ECOLI	1.17e+00
32	86	5.2	934	1	EAE2_ECOLI	1.17e+00
33	86	5.2	934	1	EAE2_ECOLI	1.17e+00
34	86	5.2	934	1	EAE2_ECOLI	1.17e+00
35	86	5.2	934	1	EAE2_ECOLI	1.17e+00
36	86	5.2	934	1	EAE2_ECOLI	1.17e+00
37	86	5.2	934	1	EAE2_ECOLI	1.17e+00
38	86	5.2	934	1	EAE2_ECOLI	1.17e+00
39	86	5.2	934	1	EAE2_ECOLI	1.17e+00
40	86	5.2	934	1	EAE2_ECOLI	1.17e+00
41	86	5.2	934	1	EAE2_ECOLI	1.17e+00
42	86	5.2	934	1	EAE2_ECOLI	1.17e+00
43	86	5.2	934	1	EAE2_ECOLI	1.17e+00
44	86	5.2	934	1	EAE2_ECOLI	1.17e+00
45	86	5.2	934	1	EAE2_ECOLI	1.17e+00

ALIGNMENTS

RESULT 1  
ID CYRG CANFA STANDARD: PRT: 373 AA.  
AC P40321;  
DT 01-FEB-1995 (rel. 31, Last sequence update)  
DT 01-FEB-1995 (rel. 31, Last sequence update)  
DT 15-JUL-1998 (rel. 36, Last annotation update)  
DE CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C)  
DE (INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64).  
GN IL2RG.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-SPLEEN;  
RX MEDLINE: 95130114.  
PA Henthorn P.S., Somborg R.L., Fimiani V.M., Puck J.M., Patterson D.F.,  
PA Falsburg P.J.;  
RT "IL-2R gamma gene microdeletion demonstrates that canine x-linked  
RT severe combined immunodeficiency is a homologue of the human  
RT disease";  
RT Genomics 23:69-74(1994).  
CC -1- FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF  
CC -1- INTERLEUKINS.  
CC -1- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND  
CC -1- PROBABLY ALSO THE IL-13 RECEPTORS.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- DISEASE: DEFECTS IN IL2RG ARE THE CAUSE OF A CANINE X-LINKED  
CC -1- SEVERE COMBINED IMMUNODEFICIENCY.  
CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC at the European Bioinformatics Institute where there are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Used by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -1- EMBL: U04361; AAC48403.1;  
CC -1- HSSP: P31785; ILIN.  
CC -1- PROSITE: PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
CC -1- PROSITE: PS00340; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.  
CC -1- PFM: PFM0041; fn3; 1.  
CC -1- SIGNAL: Transmembrane; glycoprotein; signal.  
CC -1- POTENTIAL. 22

FT CHAIN 23 373 CYTOKINE RECEPTOR COMMON GAMMA CHAIN.  
 FT DOMAIN 23 261 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 262 283 POTENTIAL.  
 FT DOMAIN 284 373 CYTOSOLIC (POTENTIAL).  
 FT DISULFID 151 249 FIBRONECTIN TYPE-III.  
 FT DISULFID 62 72 POTENTIAL.  
 FT CARBOHYD 102 113 POTENTIAL.  
 FT CARBOHYD 24 24 POTENTIAL.  
 FT CARBOHYD 71 71 POTENTIAL.  
 FT CARBOHYD 75 75 POTENTIAL.  
 FT CARBOHYD 84 84 POTENTIAL.  
 FT CARBOHYD 159 159 POTENTIAL.  
 FT CARBOHYD 164 164 POTENTIAL.  
 FT CARBOHYD 249 249 POTENTIAL.  
 SO SEQUENCE 373 AA; 42516 MW; 03A0DE1F8B089DB8 CRC64;

Query Match 13.2%; Score 210; DB 1; Length 373;  
 Best Local Similarity 25.9%; Pred No. 5 01e-24;  
 Matches 49; Conservative 56; Mismatches 73; Indels 11; Gaps 10;

Db 59 EVGFVFEVFNCTWSSSEPRPTNLTLYWYKNSNDKVOEGHYFSEVTAQWLO 118  
 QY 31 QIQLITVFNLEVOYTWASKYSR-TNLFHYRF-NGDEA-YDQCTNYLLOGHSGCLLD 87  
 Db 119 KEELHLEYEVVQLRDPREPSTOKLQNLVTPAPENLTHNSESOLEISWSNRH 178  
 QY 88 AEQRD-DILYFSIRNGTHPYFAS-RMNYTL-KPSSPKHYRF-SWHDVAVTQCDLS 142  
 Db 179 LDHCLHENVQYRSWDRSMTQSVDRNSFSLPSVDQKFTYFVRSRY-NPLGSGQRW 237  
 QY 143 YGD-LLEYVGRSEFDEW-QSKQENTCNVTIEGLDAKCSFVWRKAMEDYGPDTYP 200  
 Db 238 SEMSHPIHW 246  
 QY 201 SDWSEVTCW 209

RESULT 2  
 ID CYRG\_HUMAN STANDARD; PRT; 369 AA.  
 AC P31785;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C) (INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64) (CD132 ANTIGEN).  
 GN IL2RG.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RA Takeshita T., Asao H., Ohtani K., Ishii N., Kumaki S., Tanaka N.,  
 RA Munakata H., Nakamura M., Sugamura K.;  
 RA "Cloning of the gamma chain of the human IL-2 receptor.";  
 RA Science 257:379-382(1992).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE-LIVER.  
 RX MEDLINE: 93293887.  
 RA Noguchi M., Adelstein S., Cao X., Leonard W.J.;  
 RA "Characterization of the human interleukin-2 receptor gamma chain  
 RL gene.";  
 RL J. Biol. Chem. 268:13601-13608(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND VARIANTS ASP-114 AND ASN-153.  
 RX MEDLINE: 94004847.  
 RA Puck J.M., Deschenes S.M., Porter J.C., Dutra A.S., Brown C.J.,  
 RA Willard H., Henthorn P.S.;  
 RA "The interleukin-2 receptor gamma chain maps to Xq13.1 and is mutated  
 RL in X-linked severe combined immunodeficiency, SCIDX1.";  
 RL Hum. Mol. Genet. 2:1099-1104(1993).  
 RN [4]

RP IDENTIFICATION AS A IL-4R SUBUNIT.  
 RX MEDLINE: 94090315.  
 RA Kondo M., Takeshita T., Ishii N., Nakamura M., Watanabe S.,  
 RA Arai K.-I., Sugamura K.;  
 RA "Sharing of the interleukin-2 (IL-2) receptor gamma chain between  
 RL receptors for IL-2 and IL-4.";  
 RL Science 262:1874-1877(1993).  
 RN [5]  
 RP IDENTIFICATION AS A IL-4R SUBUNIT.  
 RX MEDLINE: 94090317.  
 RA Russell S.M., Kegan A.D., Harada N., Nakamura Y., Noguchi M.,  
 RA Leonard P., Friedmann M.C., Miyajima A., Puri R.K., Paul W.E.,  
 RA Leonard W.J.;  
 RA "Interleukin-2 receptor gamma chain: a functional component of the  
 RL interleukin-4 receptor.";  
 RL Science 262:1880-1883(1993).  
 RN [6]  
 RP IDENTIFICATION AS A IL-7R SUBUNIT.  
 RX MEDLINE: 94090316.  
 RA Noguchi M., Nakamura Y., Russell S.M., Ziegler S.F., Tsang M., Cao X.,  
 RA Leonard W.J.;  
 RA "Interleukin-2 receptor gamma chain: a functional component of the  
 RL interleukin-7 receptor.";  
 RL Science 262:1877-1880(1993).  
 RN [7]  
 RP 3D-STRUCTURE MODELING OF 57-248.  
 RX MEDLINE: 95111955.  
 RA Bamorough P., Hedgecock C.J., Richards W.G.;  
 RA "The interleukin-2 and interleukin-4 receptors studied by molecular  
 RL modelling.";  
 RL Structure 2:839-851(1994).  
 RN [8]  
 RP VARIANTS XSCID PHE-115; CYS-240 AND ILE-241.  
 RX MEDLINE: 94130970.  
 RA Disanto J.P., Dautry-Varsat A., Certain S., Fischer A.,  
 RA de Saint Basile G.;  
 RA "Interleukin-2 (IL-2) receptor gamma chain mutations in X-linked  
 RL severe combined immunodeficiency disease result in the loss of  
 RL high-affinity IL-2 receptor binding.";  
 RL Eur. J. Immunol. 24:475-479(1994).  
 RN [9]  
 RP VARIANT XSCID LYS-68.  
 RX MEDLINE: 94375038.  
 RA Marilewicz S., Subtil A., Dautry-Varsat A., Fischer A.,  
 RA de Saint Basile G.;  
 RA "Detection of three nonsense mutations and one missense mutation in  
 RL the interleukin-2 receptor gamma chain gene in SCIDX1 that  
 RL differently affect the mRNA processing.";  
 RL Genomics 21:291-293(1994).  
 RN [10]  
 RP VARIANT XSCID HIS-162.  
 RX MEDLINE: 94300093.  
 RA Ishii N., Asao H., Kimura Y., Takeshita T., Nakamura M., Tsuchiya S.,  
 RA Kono T., Maeda M., Uchiyama T., Sugamura K.;  
 RA "Impairment of ligand binding and growth signaling of mutant IL-2  
 RL receptor gamma-chains in patients with X-linked severe combined  
 RL immunodeficiency.";  
 RL J. Immunol. 153:1310-1317(1994).  
 RN [11]  
 RP VARIANT XSCID ASN-39.  
 RX MEDLINE: 95023932.  
 RA Disanto J.P., Rieux-Laucat F., Dautry-Varsat A., Fischer A.,  
 RA de Saint Basile G.;  
 RA "Defective human interleukin 2 receptor gamma chain in an atypical X  
 RL chromosome-linked severe combined immunodeficiency with peripheral T  
 RL cells.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:9466-9470(1994).  
 RN [12]  
 RP VARIANTS XSCID CYS-226 AND HIS-226.  
 RX MEDLINE: 95397841.  
 RA Pepper A.E., Buckley R.H., Small T.N., Puck J.M.;  
 RA "Two mutational hotspots in the interleukin-2 receptor gamma chain  
 RL gene causing human X-linked severe combined immunodeficiency.";



Am. J. Hum. Genet. 57:564-571(1995).

[13]  
 RN VARIANT XSCID SER-183.  
 RP MEDLINE: 96013903.  
 RA Clark P.A., Lester T., Genet S., Jones A.M., Hendriks R.,  
 Ravitsky R.L., Kinnon C.;  
 RT "Screening for mutations causing X-linked severe combined  
 immunodeficiency in the IL-2R gamma chain gene by single-strand  
 conformation polymorphism analysis.";  
 RL Hum. Genet. 96:427-432(1995).

[14]  
 RP VARIANT XSCID GLN-HIS-TRP INS-237.  
 RA MEDLINE: 95164726.  
 RT "Female germ line mosaicism as the origin of a unique IL-2 receptor  
 gamma-chain mutation causing X-linked severe combined  
 immunodeficiency.";  
 RL J. Clin. Invest. 95:895-899(1995).

[15]  
 RP VARIANT XSCID GLN-271.  
 RA MEDLINE: 95190013.  
 RT "Schmalstieg F.C., Leonard W.J., Noguichi M., Berg M., Rudloff H.E.,  
 Denney R.M., Dave S.K., Brooks E.G., Goldman A.S.;  
 RT "Missense mutation in exon 7 of the common gamma chain causes a  
 disease form of X-linked combined immunodeficiency.";  
 RL J. Clin. Invest. 95:1169-1173(1995).

[16]  
 RP VARIANT XSCID ARG-115.  
 RA MEDLINE: 97042245.  
 RT "Stephan V., Kahn V., Le Deist F., Dirksen U., Broker B.,  
 Muller-Fleckenstein I., Horneff G., Schroten H., Fischer A.,  
 de Saint Basile G.;  
 RT "Atypical X-linked severe combined immunodeficiency due to possible  
 spontaneous reversion of the genetic defect in T cells.";  
 RL New Engl. J. Med. 335:1563-1567(1996).

[17]  
 RP VARIANT XSCID GLN-285.  
 RA MEDLINE: 97295088.  
 RT "Jones A.M., Clark P.A., Katz F., Genet S., McMahon C., Alterman L.,  
 Cant A., Kinnon C.;  
 RT "B-cell-negative severe combined immunodeficiency associated with a  
 common gamma chain mutation.";  
 RL Hum. Genet. 99:677-680(1997).

[18]  
 RP VARIANT XSCID CYS-222.  
 RA MEDLINE: 98064061.  
 RT "Sharfe N., Shabar M., Roifman C.M.;  
 RT "An interleukin-2 receptor gamma chain mutation with normal thymus  
 morphology.";  
 RL J. Clin. Invest. 100:3036-3043(1997).

[19]  
 RP FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF  
 INTERLEUKINS.  
 CC -1- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND  
 IL-13 RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: TYPE 1 MEMBRANE PROTEIN.  
 CC -1- DISEASE: DEFECTS IN IL2RG ARE THE CAUSE OF A SEVERE COMBINED  
 IMMUNODEFICIENCY, WHICH IS KNOWN AS AGAMMAGLOBULINEMIA, SWISS TYPE  
 OR X-LINKED SEVERE COMBINED IMMUNODEFICIENCY DISEASE (XSCID).  
 CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD132 entry.  
 CC -1- DATABASE: NAME=IL2RGbase; NOTE=X-linked SCID mutation database;  
 CC -1- WWW="http://www.ncbi.nlm.nih.gov/row/cd/cd132.htm".  
 CC -1- WWW="http://www.ncbi.nlm.nih.gov/row/cd/cd132.htm".  
 CC -1- WWW="http://www.ncbi.nlm.nih.gov/row/cd/cd132.htm".  
 CC -1- WWW="http://www.ncbi.nlm.nih.gov/row/cd/cd132.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR DBL: D11086; BAA01857.1; -  
 DR DBL: L12183; AAA59145.1; -  
 DR DBL: L12178; AAA59145.1; JOINED.  
 DR DBL: L12176; AAA59145.1; JOINED.  
 DR DBL: L12177; AAA59145.1; JOINED.  
 DR DBL: L12179; AAA59145.1; JOINED.  
 DR DBL: L12180; AAA59145.1; JOINED.  
 DR DBL: L12181; AAA59145.1; JOINED.  
 DR DBL: L12182; AAA59145.1; JOINED.  
 DR DBL: L12183; AAA59145.1; JOINED.  
 DR DBL: L12184; AAA59145.1; JOINED.  
 DR DBL: L12185; AAA59145.1; JOINED.  
 DR DBL: L12186; AAA59145.1; JOINED.  
 DR DBL: L12187; AAA59145.1; JOINED.  
 DR DBL: L12188; AAA59145.1; JOINED.  
 DR DBL: L12189; AAA59145.1; JOINED.  
 DR DBL: L12190; AAA59145.1; JOINED.  
 DR DBL: L12191; AAA59145.1; JOINED.  
 DR DBL: L12192; AAA59145.1; JOINED.  
 DR DBL: L12193; AAA59145.1; JOINED.  
 DR DBL: L12194; AAA59145.1; JOINED.  
 DR DBL: L12195; AAA59145.1; JOINED.  
 DR DBL: L12196; AAA59145.1; JOINED.  
 DR DBL: L12197; AAA59145.1; JOINED.  
 DR DBL: L12198; AAA59145.1; JOINED.  
 DR DBL: L12199; AAA59145.1; JOINED.  
 DR DBL: L12200; AAA59145.1; JOINED.  
 DR DBL: L12201; AAA59145.1; JOINED.  
 DR DBL: L12202; AAA59145.1; JOINED.  
 DR DBL: L12203; AAA59145.1; JOINED.  
 DR DBL: L12204; AAA59145.1; JOINED.  
 DR DBL: L12205; AAA59145.1; JOINED.  
 DR DBL: L12206; AAA59145.1; JOINED.  
 DR DBL: L12207; AAA59145.1; JOINED.  
 DR DBL: L12208; AAA59145.1; JOINED.  
 DR DBL: L12209; AAA59145.1; JOINED.  
 DR DBL: L12210; AAA59145.1; JOINED.  
 DR DBL: L12211; AAA59145.1; JOINED.  
 DR DBL: L12212; AAA59145.1; JOINED.  
 DR DBL: L12213; AAA59145.1; JOINED.  
 DR DBL: L12214; AAA59145.1; JOINED.  
 DR DBL: L12215; AAA59145.1; JOINED.  
 DR DBL: L12216; AAA59145.1; JOINED.  
 DR DBL: L12217; AAA59145.1; JOINED.  
 DR DBL: L12218; AAA59145.1; JOINED.  
 DR DBL: L12219; AAA59145.1; JOINED.  
 DR DBL: L12220; AAA59145.1; JOINED.  
 DR DBL: L12221; AAA59145.1; JOINED.  
 DR DBL: L12222; AAA59145.1; JOINED.  
 DR DBL: L12223; AAA59145.1; JOINED.  
 DR DBL: L12224; AAA59145.1; JOINED.  
 DR DBL: L12225; AAA59145.1; JOINED.  
 DR DBL: L12226; AAA59145.1; JOINED.  
 DR DBL: L12227; AAA59145.1; JOINED.  
 DR DBL: L12228; AAA59145.1; JOINED.  
 DR DBL: L12229; AAA59145.1; JOINED.  
 DR DBL: L12230; AAA59145.1; JOINED.  
 DR DBL: L12231; AAA59145.1; JOINED.  
 DR DBL: L12232; AAA59145.1; JOINED.  
 DR DBL: L12233; AAA59145.1; JOINED.  
 DR DBL: L12234; AAA59145.1; JOINED.  
 DR DBL: L12235; AAA59145.1; JOINED.  
 DR DBL: L12236; AAA59145.1; JOINED.  
 DR DBL: L12237; AAA59145.1; JOINED.  
 DR DBL: L12238; AAA59145.1; JOINED.  
 DR DBL: L12239; AAA59145.1; JOINED.  
 DR DBL: L12240; AAA59145.1; JOINED.  
 DR DBL: L12241; AAA59145.1; JOINED.  
 DR DBL: L12242; AAA59145.1; JOINED.  
 DR DBL: L12243; AAA59145.1; JOINED.  
 DR DBL: L12244; AAA59145.1; JOINED.  
 DR DBL: L12245; AAA59145.1; JOINED.  
 DR DBL: L12246; AAA59145.1; JOINED.  
 DR DBL: L12247; AAA59145.1; JOINED.  
 DR DBL: L12248; AAA59145.1; JOINED.  
 DR DBL: L12249; AAA59145.1; JOINED.  
 DR DBL: L12250; AAA59145.1; JOINED.  
 DR DBL: L12251; AAA59145.1; JOINED.  
 DR DBL: L12252; AAA59145.1; JOINED.  
 DR DBL: L12253; AAA59145.1; JOINED.  
 DR DBL: L12254; AAA59145.1; JOINED.  
 DR DBL: L12255; AAA59145.1; JOINED.  
 DR DBL: L12256; AAA59145.1; JOINED.  
 DR DBL: L12257; AAA59145.1; JOINED.  
 DR DBL: L12258; AAA59145.1; JOINED.  
 DR DBL: L12259; AAA59145.1; JOINED.  
 DR DBL: L12260; AAA59145.1; JOINED.  
 DR DBL: L12261; AAA59145.1; JOINED.  
 DR DBL: L12262; AAA59145.1; JOINED.  
 DR DBL: L12263; AAA59145.1; JOINED.  
 DR DBL: L12264; AAA59145.1; JOINED.  
 DR DBL: L12265; AAA59145.1; JOINED.  
 DR DBL: L12266; AAA59145.1; JOINED.  
 DR DBL: L12267; AAA59145.1; JOINED.  
 DR DBL: L12268; AAA59145.1; JOINED.  
 DR DBL: L12269; AAA59145.1; JOINED.  
 DR DBL: L12270; AAA59145.1; JOINED.  
 DR DBL: L12271; AAA59145.1; JOINED.  
 DR DBL: L12272; AAA59145.1; JOINED.  
 DR DBL: L12273; AAA59145.1; JOINED.  
 DR DBL: L12274; AAA59145.1; JOINED.  
 DR DBL: L12275; AAA59145.1; JOINED.  
 DR DBL: L12276; AAA59145.1; JOINED.  
 DR DBL: L12277; AAA59145.1; JOINED.  
 DR DBL: L12278; AAA59145.1; JOINED.  
 DR DBL: L12279; AAA59145.1; JOINED.  
 DR DBL: L12280; AAA59145.1; JOINED.  
 DR DBL: L12281; AAA59145.1; JOINED.  
 DR DBL: L12282; AAA59145.1; JOINED.  
 DR DBL: L12283; AAA59145.1; JOINED.  
 DR DBL: L12284; AAA59145.1; JOINED.  
 DR DBL: L12285; AAA59145.1; JOINED.  
 DR DBL: L12286; AAA59145.1; JOINED.  
 DR DBL: L12287; AAA59145.1; JOINED.  
 DR DBL: L12288; AAA59145.1; JOINED.  
 DR DBL: L12289; AAA59145.1; JOINED.  
 DR DBL: L12290; AAA59145.1; JOINED.  
 DR DBL: L12291; AAA59145.1; JOINED.  
 DR DBL: L12292; AAA59145.1; JOINED.  
 DR DBL: L12293; AAA59145.1; JOINED.  
 DR DBL: L12294; AAA59145.1; JOINED.  
 DR DBL: L12295; AAA59145.1; JOINED.  
 DR DBL: L12296; AAA59145.1; JOINED.  
 DR DBL: L12297; AAA59145.1; JOINED.  
 DR DBL: L12298; AAA59145.1; JOINED.  
 DR DBL: L12299; AAA59145.1; JOINED.  
 DR DBL: L12300; AAA59145.1; JOINED.  
 DR DBL: L12301; AAA59145.1; JOINED.  
 DR DBL: L12302; AAA59145.1; JOINED.  
 DR DBL: L12303; AAA59145.1; JOINED.  
 DR DBL: L12304; AAA59145.1; JOINED.  
 DR DBL: L12305; AAA59145.1; JOINED.  
 DR DBL: L12306; AAA59145.1; JOINED.  
 DR DBL: L12307; AAA59145.1; JOINED.  
 DR DBL: L12308; AAA59145.1; JOINED.  
 DR DBL: L12309; AAA59145.1; JOINED.  
 DR DBL: L12310; AAA59145.1; JOINED.  
 DR DBL: L12311; AAA59145.1; JOINED.  
 DR DBL: L12312; AAA59145.1; JOINED.  
 DR DBL: L12313; AAA59145.1; JOINED.  
 DR DBL: L12314; AAA59145.1; JOINED.  
 DR DBL: L12315; AAA59145.1; JOINED.  
 DR DBL: L12316; AAA59145.1; JOINED.  
 DR DBL: L12317; AAA59145.1; JOINED.  
 DR DBL: L12318; AAA59145.1; JOINED.  
 DR DBL: L12319; AAA59145.1; JOINED.  
 DR DBL: L12320; AAA59145.1; JOINED.  
 DR DBL: L12321; AAA59145.1; JOINED.  
 DR DBL: L12322; AAA59145.1; JOINED.  
 DR DBL: L12323; AAA59145.1; JOINED.  
 DR DBL: L12324; AAA59145.1; JOINED.  
 DR DBL: L12325; AAA59145.1; JOINED.  
 DR DBL: L12326; AAA59145.1; JOINED.  
 DR DBL: L12327; AAA59145.1; JOINED.  
 DR DBL: L12328; AAA59145.1; JOINED.  
 DR DBL: L12329; AAA59145.1; JOINED.  
 DR DBL: L12330; AAA59145.1; JOINED.  
 DR DBL: L12331; AAA59145.1; JOINED.  
 DR DBL: L12332; AAA59145.1; JOINED.  
 DR DBL: L12333; AAA59145.1; JOINED.  
 DR DBL: L12334; AAA59145.1; JOINED.  
 DR DBL: L12335; AAA59145.1; JOINED.  
 DR DBL: L12336; AAA59145.1; JOINED.  
 DR DBL: L12337; AAA59145.1; JOINED.  
 DR DBL: L12338; AAA59145.1; JOINED.  
 DR DBL: L12339; AAA59145.1; JOINED.  
 DR DBL: L12340; AAA59145.1; JOINED.  
 DR DBL: L12341; AAA59145.1; JOINED.  
 DR DBL: L12342; AAA59145.1; JOINED.  
 DR DBL: L12343; AAA59145.1; JOINED.  
 DR DBL: L12344; AAA59145.1; JOINED.  
 DR DBL: L12345; AAA59145.1; JOINED.  
 DR DBL: L12346; AAA59145.1; JOINED.  
 DR DBL: L12347; AAA59145.1; JOINED.  
 DR DBL: L12348; AAA59145.1; JOINED.  
 DR DBL: L12349; AAA59145.1; JOINED.  
 DR DBL: L12350; AAA59145.1; JOINED.  
 DR DBL: L12351; AAA59145.1; JOINED.  
 DR DBL: L12352; AAA59145.1; JOINED.  
 DR DBL: L12353; AAA59145.1; JOINED.  
 DR DBL: L12354; AAA59145.1; JOINED.  
 DR DBL: L12355; AAA59145.1; JOINED.  
 DR DBL: L12356; AAA59145.1; JOINED.  
 DR DBL: L12357; AAA59145.1; JOINED.  
 DR DBL: L12358; AAA59145.1; JOINED.  
 DR DBL: L12359; AAA59145.1; JOINED.  
 DR DBL: L12360; AAA59145.1; JOINED.  
 DR DBL: L12361; AAA59145.1; JOINED.  
 DR DBL: L12362; AAA59145.1; JOINED.  
 DR DBL: L12363; AAA59145.1; JOINED.  
 DR DBL: L12364; AAA59145.1; JOINED.  
 DR DBL: L12365; AAA59145.1; JOINED.  
 DR DBL: L12366; AAA59145.1; JOINED.  
 DR DBL: L12367; AAA59145.1; JOINED.  
 DR DBL: L12368; AAA59145.1; JOINED.  
 DR DBL: L12369; AAA59145.1; JOINED.  
 DR DBL: L12370; AAA59145.1; JOINED.  
 DR DBL: L12371; AAA59145.1; JOINED.  
 DR DBL: L12372; AAA59145.1; JOINED.  
 DR DBL: L12373; AAA59145.1; JOINED.  
 DR DBL: L12374; AAA59145.1; JOINED.  
 DR DBL: L12375; AAA59145.1; JOINED.  
 DR DBL: L12376; AAA59145.1; JOINED.  
 DR DBL: L12377; AAA59145.1; JOINED.  
 DR DBL: L12378; AAA59145.1; JOINED.  
 DR DBL: L12379; AAA59145.1; JOINED.  
 DR DBL: L12380; AAA59145.1; JOINED.  
 DR DBL: L12381; AAA59145.1; JOINED.  
 DR DBL: L12382; AAA59145.1; JOINED.  
 DR DBL: L12383; AAA59145.1; JOINED.  
 DR DBL: L12384; AAA59145.1; JOINED.  
 DR DBL: L12385; AAA59145.1; JOINED.  
 DR DBL: L12386; AAA59145.1; JOINED.  
 DR DBL: L12387; AAA59145.1; JOINED.  
 DR DBL: L12388; AAA59145.1; JOINED.  
 DR DBL: L12389; AAA59145.1; JOINED.  
 DR DBL: L12390; AAA59145.1; JOINED.  
 DR DBL: L12391; AAA59145.1; JOINED.  
 DR DBL: L12392; AAA59145.1; JOINED.  
 DR DBL: L12393; AAA59145.1; JOINED.  
 DR DBL: L12394; AAA59145.1; JOINED.  
 DR DBL: L12395; AAA59145.1; JOINED.  
 DR DBL: L12396; AAA59145.1; JOINED.  
 DR DBL: L12397; AAA59145.1; JOINED.  
 DR DBL: L12398; AAA59145.1; JOINED.  
 DR DBL: L12399; AAA59145.1; JOINED.  
 DR DBL: L12400; AAA59145.1; JOINED.  
 DR DBL: L12401; AAA59145.1; JOINED.  
 DR DBL: L12402; AAA59145.1; JOINED.  
 DR DBL: L12403; AAA59145.1; JOINED.  
 DR DBL: L12404; AAA59145.1; JOINED.  
 DR DBL: L12405; AAA59145.1; JOINED.  
 DR DBL: L12406; AAA59145.1; JOINED.  
 DR DBL: L12407; AAA59145.1; JOINED.  
 DR DBL: L12408; AAA59145.1; JOINED.  
 DR DBL: L12409; AAA59145.1; JOINED.  
 DR DBL: L12410; AAA59145.1; JOINED.  
 DR DBL: L12411; AAA59145.1; JOINED.  
 DR DBL: L12412; AAA59145.1; JOINED.  
 DR DBL: L12413; AAA59145.1; JOINED.  
 DR DBL: L12414; AAA59145.1; JOINED.  
 DR DBL: L12415; AAA59145.1; JOINED.  
 DR DBL: L12416; AAA59145.1; JOINED.  
 DR DBL: L12417; AAA59145.1; JOINED.  
 DR DBL: L12418; AAA59145.1; JOINED.  
 DR DBL: L12419; AAA59145.1; JOINED.  
 DR DBL: L12420; AAA59145.1; JOINED.  
 DR DBL: L12421; AAA59145.1; JOINED.  
 DR DBL: L12422; AAA59145.1; JOINED.  
 DR DBL: L12423; AAA59145.1; JOINED.  
 DR DBL: L12424; AAA59145.1; JOINED.  
 DR DBL: L12425; AAA59145.1; JOINED.  
 DR DBL: L12426; AAA59145.1; JOINED.  
 DR DBL: L12427; AAA59145.1; JOINED.  
 DR DBL: L12428; AAA59145.1; JOINED.  
 DR DBL: L12429; AAA59145.1; JOINED.  
 DR DBL: L12430; AAA59145.1; JOINED.  
 DR DBL: L12431; AAA59145.1; JOINED.  
 DR DBL: L12432; AAA59145.1; JOINED.  
 DR DBL: L12433; AAA59145.1; JOINED.  
 DR DBL: L12434; AAA59145.1; JOINED.  
 DR DBL: L12435; AAA59145.1; JOINED.  
 DR DBL: L12436; AAA59145.1; JOINED.  
 DR DBL: L12437; AAA59145.1; JOINED.  
 DR DBL: L12438; AAA59145.1; JOINED.  
 DR DBL: L12439; AAA59145.1; JOINED.  
 DR DBL: L12440; AAA59145.1; JOINED.  
 DR DBL: L12441; AAA59145.1; JOINED.  
 DR DBL: L12442; AAA59145.1; JOINED.  
 DR DBL: L12443; AAA59145.1; JOINED.  
 DR DBL: L12444; AAA59145.1; JOINED.  
 DR DBL: L12445; AAA59145.1; JOINED.  
 DR DBL: L12446; AAA59145.1; JOINED.  
 DR DBL: L12447; AAA59145.1; JOINED.  
 DR DBL: L12448; AAA59145.1; JOINED.  
 DR DBL: L12449; AAA59145.1; JOINED.  
 DR DBL: L12450; AAA59145.1; JOINED.  
 DR DBL: L12451; AAA59145.1; JOINED.  
 DR DBL: L12452; AAA59145.1; JOINED.  
 DR DBL: L12453; AAA59145.1; JOINED.  
 DR DBL: L12454; AAA59145.1; JOINED.  
 DR DBL: L12455; AAA59145.1; JOINED.  
 DR DBL: L12456; AAA59145.1; JOINED.  
 DR DBL: L12457; AAA59145.1; JOINED.  
 DR DBL: L12458; AAA59145.1; JOINED.  
 DR DBL: L12459; AAA59145.1; JOINED.  
 DR DBL: L12460; AAA59145.1; JOINED.  
 DR DBL: L12461; AAA59145.1; JOINED.  
 DR DBL: L12462; AAA59145.1; JOINED.  
 DR DBL: L12463; AAA59145.1; JOINED.  
 DR DBL: L12464; AAA59145.1; JOINED.  
 DR DBL: L12465; AAA59145.1; JOINED.  
 DR DBL: L12466; AAA59145.1; JOINED.  
 DR DBL: L12467; AAA59145.1; JOINED.  
 DR DBL: L12468; AAA59145.1; JOINED.  
 DR DBL: L12469; AAA59145.1; JOINED.  
 DR DBL: L12470; AAA59145.1; JOINED.  
 DR DBL: L12471; AAA59145.1; JOINED.  
 DR DBL: L12472; AAA59145.1; JOINED.  
 DR DBL: L12473; AAA59145.1; JOINED.  
 DR DBL: L12474; AAA59145.1; JOINED.  
 DR DBL: L12475; AAA59145.1; JOINED.  
 DR DBL: L12476; AAA59145.1; JOINED.  
 DR DBL: L12477; AAA59145.1; JOINED.  
 DR DBL: L12478; AAA59145.1; JOINED.  
 DR DBL: L12479; AAA59145.1; JOINED.  
 DR DBL: L12480; AAA59145.1; JOINED.  
 DR DBL: L12481; AAA59145.1; JOINED.  
 DR DBL: L12482; AAA59145.1; JOINED.  
 DR DBL: L12483; AAA59145.1; JOINED.  
 DR DBL: L12484; AAA59145.1; JOINED.  
 DR DBL: L12485; AAA59145.1; JOINED.  
 DR DBL: L12486; AAA59145.1; JOINED.  
 DR DBL: L12487; AAA59145.1; JOINED.  
 DR DBL: L12488; AAA59145.1; JOINED.  
 DR DBL: L12489; AAA59145.1; JOINED.  
 DR DBL: L12490; AAA59145.1; JOINED.  
 DR DBL: L12491; AAA59145.1; JOINED.  
 DR DBL: L12492; AAA59145.1; JOINED.  
 DR DBL: L12493; AAA59145.1; JOINED.  
 DR DBL: L12494; AAA59145.1; JOINED.  
 DR DBL: L12495; AAA59145.1; JOINED.  
 DR DBL: L12496; AAA59145.1; JOINED.  
 DR DBL: L12497; AAA59145.1; JOINED.  
 DR DBL: L12498; AAA59145.1; JOINED.  
 DR DBL: L12499; AAA59145.1; JOINED.  
 DR DBL: L12500; AAA59145.1; JOINED.  
 DR DBL: L12501; AAA59145.1; JOINED.  
 DR DBL: L12502; AAA59145.1; JOINED.  
 DR DBL: L12503; AAA59145.1; JOINED.  
 DR DBL: L12504; AAA59145.1; JOINED.  
 DR DBL: L12505; AAA59145.1; JOINED.  
 DR DBL: L12506; AAA59145.1; JOINED.  
 DR DBL: L12507; AAA59145.1; JOINED.  
 DR DBL: L12508; AAA59145.1; JOINED.  
 DR DBL: L12509; AAA59145.1; JOINED.  
 DR DBL: L12510; AAA59145.1; JOINED.  
 DR DBL: L12511; AAA59145.1; JOINED.  
 DR DBL: L12512; AAA59145.1; JOINED.  
 DR DBL: L12513; AAA59145.1; JOINED.  
 DR DBL: L12514; AAA59145.1; JOINED.  
 DR DBL: L12515; AAA59145.1; JOINED.  
 DR DBL: L12516; AAA59145.1; JOINED.  
 DR DBL: L12517; AAA59145.1; JOINED.  
 DR DBL: L12518; AAA59145.1; JOINED.  
 DR DBL: L12519; AAA59145.1; JOINED.  
 DR DBL: L12520; AAA59145.1; JOINED.  
 DR DBL: L12521; AAA59145.1; JOINED.  
 DR DBL: L12522; AAA59145.1; JOINED.  
 DR DBL: L12523; AAA59145.1; JOINED.  
 DR DBL: L12524; AAA59145.1; JOINED.  
 DR DBL: L12525; AAA59145.1; JOINED.  
 DR DBL: L12526; AAA59145.1; JOINED.  
 DR DBL: L12527; AAA59145.1; JOINED.  
 DR DBL: L12528; AAA59145.1; JOINED.  
 DR DBL: L12529; AAA59145.1; JOINED.  
 DR DBL: L12530; AAA59145.1; JOINED.  
 DR DBL: L12531; AAA59145.1; JOINED.  
 DR DBL: L12532; AAA59145.1; JOINED.  
 DR DBL: L12533; AAA59145.1; JOINED.  
 DR DBL: L12534; AAA59145.1; JOINED.  
 DR DBL: L12535; AAA59145.1; JOINED.  
 DR DBL: L12536; AAA59145.1; JOINED.  
 DR DBL: L12537; AAA59145.1; JOINED.  
 DR DBL: L12538; AAA59145.1; JOINED.  
 DR DBL: L12539; AAA59145.1; JOINED.  
 DR DBL: L12540; AAA59145.1; JOINED.  
 DR DBL: L12541; AAA59145.1; JOINED.  
 DR DBL: L12542; AAA59145.1; JOINED.  
 DR DBL: L12543; AAA59145.1; JOINED.  
 DR DBL: L12544; AAA59145.1; JOINED.  
 DR DBL: L12545; AAA59145.1; JOINED.  
 DR DBL: L12546; AAA59145.1; JOINED.  
 DR DBL: L12547; AAA59145.1; JOINED.  
 DR DBL: L12548; AAA59145.1; JOINED.  
 DR DBL: L12549; AAA59145.1; JOINED.  
 DR DBL: L12550; AAA59145.1; JOINED.  
 DR DBL: L12551; AAA59145.1; JOINED.  
 DR DBL: L12552; AAA59145.1; JOINED.  
 DR DBL: L12553; AAA59145.1; JOINED.  
 DR DBL: L12554; AAA59145.1; JOINED.  
 DR DBL: L12555; AAA59145.1; JOINED.  
 DR DBL: L12556; AAA59145.1; JOINED.  
 DR DBL: L12557; AAA59145.1; JOINED.  
 DR DBL: L12558; AAA59145.1; JOINED.  
 DR DBL: L12559; AAA59145.1; JOINED.  
 DR DBL: L12560; AAA59145.1; JOINED.  
 DR DBL: L12561; AAA59145.1; JOINED.  
 DR DBL: L12562; AAA59145.1; JOINED.  
 DR DBL: L12563; AAA59145.1; JOINED.  
 DR DBL: L12564; AAA59145.1; JOINED.  
 DR DBL: L12565; AAA59145.1; JOINED.  
 DR DBL: L12566; AAA59145.1; JOINED.  
 DR DBL: L12567; AAA59145.1; JOINED.  
 DR DBL: L12568; AAA59145.1; JOINED.  
 DR DBL: L12569; AAA59145.1; JOINED.  
 DR DBL: L12570; AAA59145.1; JOINED.  
 DR DBL: L12571; AAA59145.1; JOINED.<







Thu May 11 06:49:33 2000

US-09-376-430-2-04.DSP

Page 7

RA MEDLINE: 93278381.  
RA Amselem S., Duquesnoy P., Duriez B., Dastot F., Sorbier M.-L.,  
RA Vallet S., Goossens M., Duriez B., Dastot F., Sorbier M.-L.,  
RA "Spectrum of growth hormone receptor mutations and associated  
RT haplotypes in Laron syndrome."  
RT Hum. Mol. Genet. 2:355-359(1993).  
RL [6]  
RN VARIANT LARON DWARFISM HIS-170.  
RP MEDLINE: 94185645.  
RP Duquesnoy P., Sorbier M.-L., Duriez B., Dastot F., Buchanan C.R.,  
RA Savage M.O., Preece M.A., Craescu C.T., Blouquit Y., Goossens M.,  
RA Amselem S., Duriez B., Dastot F., Sorbier M.-L., Duriez B.,  
RA "A single amino acid substitution in the extracellular domain of the  
RT human growth hormone (GH) receptor confers familial GH resistance  
RT (Laron syndrome) with positive GH-binding activity by abolishing  
RT receptor homodimerization."  
RT EMBO J. 13:1386-1395(1994).  
[7]  
RN VARIANTS IDIOPATHIC SHORT STATURE LYS-62; CYS-179 AND ASP-242.  
RP MEDLINE: 96013502.  
RP Goddard A.D., Covello R., Luch S.-M., Jackson T., Attie K.M.,  
RA Gesundheit N., Rundle A.C., Wells J.A., Carlsson L.M.S.,  
RT "Mutations of the growth hormone receptor in children with idiopathic  
RT short stature. The Growth Hormone Insensitivity Study Group."  
RT New Engl. J. Med. 333:1093-1098(1995).  
[8]  
RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 19-256.  
RP MEDLINE: 92196577.  
RP de Vos A.M., Ullrich M., Kossiakoff A.A.,  
RA "Human growth hormone and extracellular domain of its receptor:  
RT crystal structure of the complex."  
RT Science 255:306-312(1992).  
[9]  
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 19-256.  
RP MEDLINE: 97113023.  
RP Sundstrem M., Lundqvist T., Rodin J., Glebel L.B., Milligan D.,  
RA Norstedt G.,  
RA "Crystal structure of an antagonist mutant of human growth hormone,  
RT G120R, in complex with its receptor at 2.9-A resolution."  
RT J. Biol. Chem. 271:32197-32203(1996).  
[10]  
RN FUNCTION: THIS IS A RECEPTOR FOR PITUITARY GLAND GROWTH HORMONE.  
CC -1- SUBUNIT: HOMODIMER.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- DISEASE: DEFICIENCY IN GHR IS THE CAUSE OF PITUITARY DWARFISM IT  
CC (LARON-TYPE PITUITARY DWARFISM OR LARON SYNDROME (LS)). IT ALSO  
CC CAUSES IDIOPATHIC SHORT STATURE.  
CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X06562; CAA29808.1; -  
CC EMBL: M28456; AAA52555.1; -  
CC EMBL: M28458; AAA52555.1; JOINED.  
CC EMBL: M28459; AAA52555.1; JOINED.  
CC EMBL: M28460; AAA52555.1; JOINED.  
CC EMBL: M28461; AAA52555.1; JOINED.  
CC EMBL: M28462; AAA52555.1; JOINED.  
CC EMBL: M28463; AAA52555.1; JOINED.  
CC EMBL: M28464; AAA52555.1; JOINED.  
CC EMBL: M28465; AAA52555.1; JOINED.  
CC EMBL: M28466; AAA52555.1; JOINED.  
CC EMBL: M28467; AAA52555.1; JOINED.  
CC EMBL: M28468; AAA52555.1; JOINED.  
CC EMBL: M28469; AAA52555.1; JOINED.  
CC EMBL: M28470; AAA52555.1; JOINED.  
CC EMBL: M28471; AAA52555.1; JOINED.  
CC EMBL: M28472; AAA52555.1; JOINED.  
CC EMBL: M28473; AAA52555.1; JOINED.  
CC EMBL: M28474; AAA52555.1; JOINED.  
CC EMBL: M28475; AAA52555.1; JOINED.  
CC EMBL: M28476; AAA52555.1; JOINED.  
CC EMBL: M28477; AAA52555.1; JOINED.  
CC EMBL: M28478; AAA52555.1; JOINED.  
CC EMBL: M28479; AAA52555.1; JOINED.  
CC EMBL: M28480; AAA52555.1; JOINED.  
CC EMBL: M28481; AAA52555.1; JOINED.  
CC EMBL: M28482; AAA52555.1; JOINED.  
CC EMBL: M28483; AAA52555.1; JOINED.  
CC EMBL: M28484; AAA52555.1; JOINED.  
CC EMBL: M28485; AAA52555.1; JOINED.  
CC EMBL: M28486; AAA52555.1; JOINED.  
CC EMBL: M28487; AAA52555.1; JOINED.  
CC EMBL: M28488; AAA52555.1; JOINED.  
CC EMBL: M28489; AAA52555.1; JOINED.  
CC EMBL: M28490; AAA52555.1; JOINED.  
CC EMBL: M28491; AAA52555.1; JOINED.  
CC EMBL: M28492; AAA52555.1; JOINED.  
CC EMBL: M28493; AAA52555.1; JOINED.  
CC EMBL: M28494; AAA52555.1; JOINED.  
CC EMBL: M28495; AAA52555.1; JOINED.  
CC EMBL: M28496; AAA52555.1; JOINED.  
CC EMBL: M28497; AAA52555.1; JOINED.  
CC EMBL: M28498; AAA52555.1; JOINED.  
CC EMBL: M28499; AAA52555.1; JOINED.  
CC EMBL: M28500; AAA52555.1; JOINED.  
CC EMBL: M28501; AAA52555.1; JOINED.  
CC EMBL: M28502; AAA52555.1; JOINED.  
CC EMBL: M28503; AAA52555.1; JOINED.  
CC EMBL: M28504; AAA52555.1; JOINED.  
CC EMBL: M28505; AAA52555.1; JOINED.  
CC EMBL: M28506; AAA52555.1; JOINED.  
CC EMBL: M28507; AAA52555.1; JOINED.  
CC EMBL: M28508; AAA52555.1; JOINED.  
CC EMBL: M28509; AAA52555.1; JOINED.  
CC EMBL: M28510; AAA52555.1; JOINED.  
CC EMBL: M28511; AAA52555.1; JOINED.  
CC EMBL: M28512; AAA52555.1; JOINED.  
CC EMBL: M28513; AAA52555.1; JOINED.  
CC EMBL: M28514; AAA52555.1; JOINED.  
CC EMBL: M28515; AAA52555.1; JOINED.  
CC EMBL: M28516; AAA52555.1; JOINED.  
CC EMBL: M28517; AAA52555.1; JOINED.  
CC EMBL: M28518; AAA52555.1; JOINED.  
CC EMBL: M28519; AAA52555.1; JOINED.  
CC EMBL: M28520; AAA52555.1; JOINED.  
CC EMBL: M28521; AAA52555.1; JOINED.  
CC EMBL: M28522; AAA52555.1; JOINED.  
CC EMBL: M28523; AAA52555.1; JOINED.  
CC EMBL: M28524; AAA52555.1; JOINED.  
CC EMBL: M28525; AAA52555.1; JOINED.  
CC EMBL: M28526; AAA52555.1; JOINED.  
CC EMBL: M28527; AAA52555.1; JOINED.  
CC EMBL: M28528; AAA52555.1; JOINED.  
CC EMBL: M28529; AAA52555.1; JOINED.  
CC EMBL: M28530; AAA52555.1; JOINED.  
CC EMBL: M28531; AAA52555.1; JOINED.  
CC EMBL: M28532; AAA52555.1; JOINED.  
CC EMBL: M28533; AAA52555.1; JOINED.  
CC EMBL: M28534; AAA52555.1; JOINED.  
CC EMBL: M28535; AAA52555.1; JOINED.  
CC EMBL: M28536; AAA52555.1; JOINED.  
CC EMBL: M28537; AAA52555.1; JOINED.  
CC EMBL: M28538; AAA52555.1; JOINED.  
CC EMBL: M28539; AAA52555.1; JOINED.  
CC EMBL: M28540; AAA52555.1; JOINED.  
CC EMBL: M28541; AAA52555.1; JOINED.  
CC EMBL: M28542; AAA52555.1; JOINED.  
CC EMBL: M28543; AAA52555.1; JOINED.  
CC EMBL: M28544; AAA52555.1; JOINED.  
CC EMBL: M28545; AAA52555.1; JOINED.  
CC EMBL: M28546; AAA52555.1; JOINED.  
CC EMBL: M28547; AAA52555.1; JOINED.  
CC EMBL: M28548; AAA52555.1; JOINED.  
CC EMBL: M28549; AAA52555.1; JOINED.  
CC EMBL: M28550; AAA52555.1; JOINED.  
CC EMBL: M28551; AAA52555.1; JOINED.  
CC EMBL: M28552; AAA52555.1; JOINED.  
CC EMBL: M28553; AAA52555.1; JOINED.  
CC EMBL: M28554; AAA52555.1; JOINED.  
CC EMBL: M28555; AAA52555.1; JOINED.  
CC EMBL: M28556; AAA52555.1; JOINED.  
CC EMBL: M28557; AAA52555.1; JOINED.  
CC EMBL: M28558; AAA52555.1; JOINED.  
CC EMBL: M28559; AAA52555.1; JOINED.  
CC EMBL: M28560; AAA52555.1; JOINED.  
CC EMBL: M28561; AAA52555.1; JOINED.  
CC EMBL: M28562; AAA52555.1; JOINED.  
CC EMBL: M28563; AAA52555.1; JOINED.  
CC EMBL: M28564; AAA52555.1; JOINED.  
CC EMBL: M28565; AAA52555.1; JOINED.  
CC EMBL: M28566; AAA52555.1; JOINED.  
CC EMBL: M28567; AAA52555.1; JOINED.  
CC EMBL: M28568; AAA52555.1; JOINED.  
CC EMBL: M28569; AAA52555.1; JOINED.  
CC EMBL: M28570; AAA52555.1; JOINED.  
CC EMBL: M28571; AAA52555.1; JOINED.  
CC EMBL: M28572; AAA52555.1; JOINED.  
CC EMBL: M28573; AAA52555.1; JOINED.  
CC EMBL: M28574; AAA52555.1; JOINED.  
CC EMBL: M28575; AAA52555.1; JOINED.  
CC EMBL: M28576; AAA52555.1; JOINED.  
CC EMBL: M28577; AAA52555.1; JOINED.  
CC EMBL: M28578; AAA52555.1; JOINED.  
CC EMBL: M28579; AAA52555.1; JOINED.  
CC EMBL: M28580; AAA52555.1; JOINED.  
CC EMBL: M28581; AAA52555.1; JOINED.  
CC EMBL: M28582; AAA52555.1; JOINED.  
CC EMBL: M28583; AAA52555.1; JOINED.  
CC EMBL: M28584; AAA52555.1; JOINED.  
CC EMBL: M28585; AAA52555.1; JOINED.  
CC EMBL: M28586; AAA52555.1; JOINED.  
CC EMBL: M28587; AAA52555.1; JOINED.  
CC EMBL: M28588; AAA52555.1; JOINED.  
CC EMBL: M28589; AAA52555.1; JOINED.  
CC EMBL: M28590; AAA52555.1; JOINED.  
CC EMBL: M28591; AAA52555.1; JOINED.  
CC EMBL: M28592; AAA52555.1; JOINED.  
CC EMBL: M28593; AAA52555.1; JOINED.  
CC EMBL: M28594; AAA52555.1; JOINED.  
CC EMBL: M28595; AAA52555.1; JOINED.  
CC EMBL: M28596; AAA52555.1; JOINED.  
CC EMBL: M28597; AAA52555.1; JOINED.  
CC EMBL: M28598; AAA52555.1; JOINED.  
CC EMBL: M28599; AAA52555.1; JOINED.  
CC EMBL: M28600; AAA52555.1; JOINED.  
CC EMBL: M28601; AAA52555.1; JOINED.  
CC EMBL: M28602; AAA52555.1; JOINED.  
CC EMBL: M28603; AAA52555.1; JOINED.  
CC EMBL: M28604; AAA52555.1; JOINED.  
CC EMBL: M28605; AAA52555.1; JOINED.  
CC EMBL: M28606; AAA52555.1; JOINED.  
CC EMBL: M28607; AAA52555.1; JOINED.  
CC EMBL: M28608; AAA52555.1; JOINED.  
CC EMBL: M28609; AAA52555.1; JOINED.  
CC EMBL: M28610; AAA52555.1; JOINED.  
CC EMBL: M28611; AAA52555.1; JOINED.  
CC EMBL: M28612; AAA52555.1; JOINED.  
CC EMBL: M28613; AAA52555.1; JOINED.  
CC EMBL: M28614; AAA52555.1; JOINED.  
CC EMBL: M28615; AAA52555.1; JOINED.  
CC EMBL: M28616; AAA52555.1; JOINED.  
CC EMBL: M28617; AAA52555.1; JOINED.  
CC EMBL: M28618; AAA52555.1; JOINED.  
CC EMBL: M28619; AAA52555.1; JOINED.  
CC EMBL: M28620; AAA52555.1; JOINED.  
CC EMBL: M28621; AAA52555.1; JOINED.  
CC EMBL: M28622; AAA52555.1; JOINED.  
CC EMBL: M28623; AAA52555.1; JOINED.  
CC EMBL: M28624; AAA52555.1; JOINED.  
CC EMBL: M28625; AAA52555.1; JOINED.  
CC EMBL: M28626; AAA52555.1; JOINED.  
CC EMBL: M28627; AAA52555.1; JOINED.  
CC EMBL: M28628; AAA52555.1; JOINED.  
CC EMBL: M28629; AAA52555.1; JOINED.  
CC EMBL: M28630; AAA52555.1; JOINED.  
CC EMBL: M28631; AAA52555.1; JOINED.  
CC EMBL: M28632; AAA52555.1; JOINED.  
CC EMBL: M28633; AAA52555.1; JOINED.  
CC EMBL: M28634; AAA52555.1; JOINED.  
CC EMBL: M28635; AAA52555.1; JOINED.  
CC EMBL: M28636; AAA52555.1; JOINED.  
CC EMBL: M28637; AAA52555.1; JOINED.  
CC EMBL: M28638; AAA52555.1; JOINED.  
CC EMBL: M28639; AAA52555.1; JOINED.  
CC EMBL: M28640; AAA52555.1; JOINED.  
CC EMBL: M28641; AAA52555.1; JOINED.  
CC EMBL: M28642; AAA52555.1; JOINED.  
CC EMBL: M28643; AAA52555.1; JOINED.  
CC EMBL: M28644; AAA52555.1; JOINED.  
CC EMBL: M28645; AAA52555.1; JOINED.  
CC EMBL: M28646; AAA52555.1; JOINED.  
CC EMBL: M28647; AAA52555.1; JOINED.  
CC EMBL: M28648; AAA52555.1; JOINED.  
CC EMBL: M28649; AAA52555.1; JOINED.  
CC EMBL: M28650; AAA52555.1; JOINED.  
CC EMBL: M28651; AAA52555.1; JOINED.  
CC EMBL: M28652; AAA52555.1; JOINED.  
CC EMBL: M28653; AAA52555.1; JOINED.  
CC EMBL: M28654; AAA52555.1; JOINED.  
CC EMBL: M28655; AAA52555.1; JOINED.  
CC EMBL: M28656; AAA52555.1; JOINED.  
CC EMBL: M28657; AAA52555.1; JOINED.  
CC EMBL: M28658; AAA52555.1; JOINED.  
CC EMBL: M28659; AAA52555.1; JOINED.  
CC EMBL: M28660; AAA52555.1; JOINED.  
CC EMBL: M28661; AAA52555.1; JOINED.  
CC EMBL: M28662; AAA52555.1; JOINED.  
CC EMBL: M28663; AAA52555.1; JOINED.  
CC EMBL: M28664; AAA52555.1; JOINED.  
CC EMBL: M28665; AAA52555.1; JOINED.  
CC EMBL: M28666; AAA52555.1; JOINED.  
CC EMBL: M28667; AAA52555.1; JOINED.  
CC EMBL: M28668; AAA52555.1; JOINED.  
CC EMBL: M28669; AAA52555.1; JOINED.  
CC EMBL: M28670; AAA52555.1; JOINED.  
CC EMBL: M28671; AAA52555.1; JOINED.  
CC EMBL: M28672; AAA52555.1; JOINED.  
CC EMBL: M28673; AAA52555.1; JOINED.  
CC EMBL: M28674; AAA52555.1; JOINED.  
CC EMBL: M28675; AAA52555.1; JOINED.  
CC EMBL: M28676; AAA52555.1; JOINED.  
CC EMBL: M28677; AAA52555.1; JOINED.  
CC EMBL: M28678; AAA52555.1; JOINED.  
CC EMBL: M28679; AAA52555.1; JOINED.  
CC EMBL: M28680; AAA52555.1; JOINED.  
CC EMBL: M28681; AAA52555.1; JOINED.  
CC EMBL: M28682; AAA52555.1; JOINED.  
CC EMBL: M28683; AAA52555.1; JOINED.  
CC EMBL: M28684; AAA52555.1; JOINED.  
CC EMBL: M28685; AAA52555.1; JOINED.  
CC EMBL: M28686; AAA52555.1; JOINED.  
CC EMBL: M28687; AAA52555.1; JOINED.  
CC EMBL: M28688; AAA52555.1; JOINED.  
CC EMBL: M28689; AAA52555.1; JOINED.  
CC EMBL: M28690; AAA52555.1; JOINED.  
CC EMBL: M28691; AAA52555.1; JOINED.  
CC EMBL: M28692; AAA52555.1; JOINED.  
CC EMBL: M28693; AAA52555.1; JOINED.  
CC EMBL: M28694; AAA52555.1; JOINED.  
CC EMBL: M28695; AAA52555.1; JOINED.  
CC EMBL: M28696; AAA52555.1; JOINED.  
CC EMBL: M28697; AAA52555.1; JOINED.  
CC EMBL: M28698; AAA52555.1; JOINED.  
CC EMBL: M28699; AAA52555.1; JOINED.  
CC EMBL: M28700; AAA52555.1; JOINED.  
CC EMBL: M28701; AAA52555.1; JOINED.  
CC EMBL: M28702; AAA52555.1; JOINED.  
CC EMBL: M28703; AAA52555.1; JOINED.  
CC EMBL: M28704; AAA52555.1; JOINED.  
CC EMBL: M28705; AAA52555.1; JOINED.  
CC EMBL: M28706; AAA52555.1; JOINED.  
CC EMBL: M28707; AAA52555.1; JOINED.  
CC EMBL: M28708; AAA52555.1; JOINED.  
CC EMBL: M28709; AAA52555.1; JOINED.  
CC EMBL: M28710; AAA52555.1; JOINED.  
CC EMBL: M28711; AAA52555.1; JOINED.  
CC EMBL: M28712; AAA52555.1; JOINED.  
CC EMBL: M28713; AAA52555.1; JOINED.  
CC EMBL: M28714; AAA52555.1; JOINED.  
CC EMBL: M28715; AAA52555.1; JOINED.  
CC EMBL: M28716; AAA52555.1; JOINED.  
CC EMBL: M28717; AAA52555.1; JOINED.  
CC EMBL: M28718; AAA52555.1; JOINED.  
CC EMBL: M28719; AAA52555.1; JOINED.  
CC EMBL: M28720; AAA52555.1; JOINED.  
CC EMBL: M28721; AAA52555.1; JOINED.  
CC EMBL: M28722; AAA52555.1; JOINED.  
CC EMBL: M28723; AAA52555.1; JOINED.  
CC EMBL: M28724; AAA52555.1; JOINED.  
CC EMBL: M28725; AAA52555.1; JOINED.  
CC EMBL: M28726; AAA52555.1; JOINED.  
CC EMBL: M28727; AAA52555.1; JOINED.  
CC EMBL: M28728; AAA52555.1; JOINED.  
CC EMBL: M28729; AAA52555.1; JOINED.  
CC EMBL: M28730; AAA52555.1; JOINED.  
CC EMBL: M28731; AAA52555.1; JOINED.  
CC EMBL: M28732; AAA52555.1; JOINED.  
CC EMBL: M28733; AAA52555.1; JOINED.  
CC EMBL: M28734; AAA52555.1; JOINED.  
CC EMBL: M28735; AAA52555.1; JOINED.  
CC EMBL: M28736; AAA52555.1; JOINED.  
CC EMBL: M28737; AAA52555.1; JOINED.  
CC EMBL: M28738; AAA52555.1; JOINED.  
CC EMBL: M28739; AAA52555.1; JOINED.  
CC EMBL: M28740; AAA52555.1; JOINED.  
CC EMBL: M28741; AAA52555.1; JOINED.  
CC EMBL: M28742; AAA52555.1; JOINED.  
CC EMBL: M28743; AAA52555.1; JOINED.  
CC EMBL: M28744; AAA52555.1; JOINED.  
CC EMBL: M28745; AAA52555.1; JOINED.  
CC EMBL: M28746; AAA52555.1; JOINED.  
CC EMBL: M28747; AAA52555.1; JOINED.  
CC EMBL: M28748; AAA52555.1; JOINED.  
CC EMBL: M28749; AAA52555.1; JOINED.  
CC EMBL: M28750; AAA52555.1; JOINED.  
CC EMBL: M28751; AAA52555.1; JOINED.  
CC EMBL: M28752; AAA52555.1; JOINED.  
CC EMBL: M28753; AAA52555.1; JOINED.  
CC EMBL: M28754; AAA52555.1; JOINED.  
CC EMBL: M28755; AAA52555.1; JOINED.  
CC EMBL: M28756; AAA52555.1; JOINED.  
CC EMBL: M28757; AAA52555.1; JOINED.  
CC EMBL: M28758; AAA52555.1; JOINED.  
CC EMBL: M28759; AAA52555.1; JOINED.  
CC EMBL: M28760; AAA52555.1; JOINED.  
CC EMBL: M28761; AAA52555.1; JOINED.  
CC EMBL: M28762; AAA52555.1; JOINED.  
CC EMBL: M28763; AAA52555.1; JOINED.  
CC EMBL: M28764; AAA52555.1; JOINED.  
CC EMBL: M28765; AAA52555.1; JOINED.  
CC EMBL: M28766; AAA52555.1; JOINED.  
CC EMBL: M28767; AAA52555.1; JOINED.  
CC EMBL: M28768; AAA52555.1; JOINED.  
CC EMBL: M28769; AAA52555.1; JOINED.  
CC EMBL: M28770; AAA52555.1; JOINED.  
CC EMBL: M28771; AAA52555.1; JOINED.  
CC EMBL: M28772; AAA52555.1; JOINED.  
CC EMBL: M28773; AAA52555.1; JOINED.  
CC EMBL: M28774; AAA52555.1; JOINED.  
CC EMBL: M28775; AAA52555.1; JOINED.  
CC EMBL: M28776; AAA52555.1; JOINED.  
CC EMBL: M28777; AAA52555.1; JOINED.  
CC EMBL: M28778; AAA52555.1; JOINED.  
CC EMBL: M28779; AAA52555.1; JOINED.  
CC EMBL: M28780; AAA52555.1; JOINED.  
CC EMBL: M28781; AAA52555.1; JOINED.  
CC EMBL: M28782; AAA52555.1; JOINED.  
CC EMBL: M28783; AAA52555.1; JOINED.  
CC EMBL: M28784; AAA52555.1; JOINED.  
CC EMBL: M28785; AAA52555.1; JOINED.  
CC EMBL: M28786; AAA52555.1; JOINED.  
CC EMBL: M28787; AAA52555.1; JOINED.  
CC EMBL: M28788; AAA52555.1; JOINED.  
CC EMBL: M28789; AAA52555.1; JOINED.  
CC EMBL: M28790; AAA52555.1; JOINED.  
CC EMBL: M28791; AAA52555.1; JOINED.  
CC EMBL: M28792; AAA52555.1; JOINED.  
CC EMBL: M28793; AAA52555.1; JOINED.  
CC EMBL: M28794; AAA52555.1; JOINED.  
CC EMBL: M28795; AAA52555.1; JOINED.  
CC EMBL: M28796; AAA52555.1; JOINED.  
CC EMBL: M28797; AAA52555.1; JOINED.  
CC EMBL: M28798; AAA52555.1; JOINED.  
CC EMBL: M28799; AAA52555.1; JOINED.  
CC EMBL: M28800; AAA52555.1; JOINED.  
CC EMBL: M28801; AAA52555.1; JOINED.  
CC EMBL: M28802; AAA52555.1; JOINED.  
CC EMBL: M28803; AAA52555.1; JOINED.  
CC EMBL: M28804; AAA52555.1; JOINED.  
CC EMBL: M28805; AAA52555.1; JOINED.  
CC EMBL: M28806; AAA52555.1; JOINED.  
CC EMBL: M28807; AAA52555.1; JOINED.  
CC EMBL: M28808; AAA52555.1; JOINED.  
CC EMBL: M28809; AAA52555.1; JOINED.  
CC EMBL: M28810; AAA52555.1; JOINED.  
CC EMBL: M28811; AAA52555.1; JOINED.  
CC EMBL: M28812; AAA52555.1; JOINED.  
CC EMBL: M28813; AAA52555.1; JOINED.  
CC EMBL: M28814; AAA52555.1; JOINED.  
CC EMBL: M28815; AAA52555.1; JOINED.  
CC EMBL: M28816; AAA52555.1; JOINED.  
CC EMBL: M28817; AAA52555.1; JOINED.  
CC EMBL: M28818; AAA52555.1; JOINED.  
CC EMBL: M28819; AAA52555.1; JOINED.  
CC EMBL: M28820; AAA52555.1; JOINED.  
CC EMBL: M28821; AAA52555.1; JOINED.  
CC EMBL: M28822; AAA52555.1; JOINED.  
CC EMBL: M28823; AAA52555.1; JOINED.  
CC EMBL: M28824; AAA52555.1; JOINED.  
CC EMBL: M28825; AAA52555.1; JOINED.  
CC EMBL: M28826; AAA52555.1; JOINED.  
CC EMBL: M28827; AAA52555.1; JOINED.  
CC EMBL: M28828; AAA52555.1; JOINED.  
CC EMBL: M28829; AAA52555.1; JOINED.  
CC EMBL: M28830; AAA52555.1; JOINED.  
CC EMBL: M28831; AAA52555.1; JOINED.  
CC EMBL: M28832; AAA52555.1; JOINED.  
CC EMBL: M28833; AAA52555.1; JOINED.  
CC EMBL: M28834; AAA52555.1; JOINED.  
CC EMBL: M28835; AAA52555.1; JOINED.  
CC EMBL: M28836; AAA52555.1; JOINED.  
CC EMBL: M28837; AAA52555.1; JOINED.  
CC EMBL: M28838; AAA52555.1; JOINED.  
CC EMBL: M28839; AAA52555.1; JOINED.  
CC EMBL: M28840; AAA52555.1; JOINED.  
CC EMBL: M28841; AAA52555.1; JOINED.  
CC EMBL: M28842; AAA52555.1; JOINED.  
CC EMBL: M28843; AAA52555.1; JOINED.  
CC EMBL: M28844; AAA52555.1; JOINED.  
CC EMBL: M28845; AAA52555.1; JOINED.  
CC EMBL: M28846; AAA52555.1; JOINED.  
CC EMBL: M28847; AAA52555.1; JOINED.  
CC EMBL: M28848; AAA52555.1; JOINED.  
CC EMBL: M28849; AAA52555.1; JOINED.  
CC EMBL: M28850; AAA52555.1; JOINED.  
CC EMBL: M28851; AAA52555.1; JOINED.  
CC EMBL: M28852; AAA52555.1; JOINED.  
CC EMBL: M28853; AAA52555.1; JOINED.  
CC EMBL: M28854; AAA52555.1; JOINED.  
CC EMBL: M28855; AAA52555.1; JOINED.  
CC EMBL: M28856; AAA52555.1; JOINED.  
CC EMBL: M28857; AAA52555.1; JOINED.  
CC EMBL: M28858; AAA52555.1; JOINED.  
CC EMBL: M28859; AAA52555.1; JOINED.  
CC EMBL: M28860; AAA52555.1; JOINED.  
CC EMBL: M28861; AAA52555.1; JOINED.  
CC EMBL: M28862; AAA52555.1; JOINED.  
CC EMBL: M28863; AAA52555.1; JOINED.  
CC EMBL: M28864; AAA52555.1; JOINED.  
CC EMBL: M28865; AAA52555.1; JOINED.  
CC EMBL: M28866; AAA52555.1; JOINED.  
CC EMBL: M28867; AAA52555.1; JOINED.  
CC EMBL: M28868; AAA52555.1; JOINED.  
CC EMBL: M28869; AAA52555.1; JOINED.  
CC EMBL: M28870; AAA52555.1; JOINED.  
CC EMBL: M28871; AAA52555.1; JOINED.  
CC EMBL: M28872; AAA52555.1; JOINED.  
CC EMBL: M28873; AAA52555.1; JOINED.  
CC EMBL: M28874; AAA52555.1; JOINED.  
CC EMBL: M28875; AAA52555.1; JOINED.  
CC EMBL: M28876; AAA52555.1; JOINED.  
CC EMBL: M28877; AAA52555.1; JOINED.  
CC EMBL: M28878; AAA52555.1; JOINED.  
CC EMBL: M28879; AAA52555.1; JOINED.  
CC EMBL: M28880; AAA52555.1; JOINED.  
CC EMBL: M28881; AAA52555.1; JOINED.  
CC EMBL: M2888

Matches 17; Conservative 11; Mismatches 22; Indels 2; Gaps 2;

Db 190 LEVELQKVENETKMKMDPILFTSPVYSLKDKKEVAVRSKORNSGYG 241  
 QY 146 LLEYVQYRSPFDEWQSKQEN-TCNVTIEGLDAKCYSFVWVKAMEDYGP 195

## RESULT 9

ID YEO6 YEAST STANDARD: PRT: 302 AA.

AC 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HYPOPHYSAL 33.5 KD PROTEIN IN PTP3-ILV1 INTERGENIC REGION  
 DE PRECURSOR.  
 GN YEO076C

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomyces.

RA [1]

RC SEQUENCE FROM N.A.

RA STRAIN-S288C / AB972;

RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,

RA Aviles E., Berro A., Brennan T., Carpenter J., Chen E., Cherry J.M.,

RA Chung E., Duncan M., Guzman E., Hartzell G., Hunkle-Smith S.,

RA Hyman R., Kayser A., Komp C., Laskarid D., Lew H., Lin D.,

RA Mosedale D., Nakahara K., Namath A., Norgren R., Oetner P., Oh C.,

RA Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,

RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.,

RA Submitted (DEC-1994) to the EMBL/GenBank/DBD databases.

CC -1 SIMILARITY: TO YEAST KILLER TOXIN KHR.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL, U18839; AAB64631.1;

DR Hypothetical protein; glycoprotein; signal.

FT SIGNAL 1 23

FT CHAIN 24 302

FT CARBOHYD 65 65

FT CARBOHYD 86 86

FT CARBOHYD 93 93

FT CARBOHYD 220 220

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

FT CARBOHYD 231 231

OS Sus scrofa (pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LANDRACE-YORKSHIRE; TISSUE-LIVER;  
 RX MEDLINE: 91057155.  
 RA Cioffi J.A., Mang X., Kopchick J.J.,  
 RT "Porcine growth hormone receptor cDNA sequence."  
 RL Nucleic Acids Res. 18:6451-6451(1990).  
 CC -1 FUNCTION: THIS IS A RECEPTOR FOR PITUITARY GLAND GROWTH HORMONE.  
 CC -1 SUBUNIT: HOMODIMER.  
 CC -1 SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1 SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -1 SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL, X54429; CA38301.1;

DR PIR, S12136; S12136.

DR HSSP, P10912; 1A22.

DR PROSITE, PS00340; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.

DR PROSITE, PS00241; RECEPTOR\_CYTOKINES\_1;

DR PFAM, PF00041; f03; 1.

KW Receptor; Transmembrane; Glycoprotein; signal.

FT SIGNAL 1 18

FT CHAIN 19 638

FT DOMAIN 19 264

FT DOMAIN 265 288

FT DOMAIN 289 338

FT DOMAIN 339 352

FT DISULFID 56 66

FT DISULFID 101 112

FT DISULFID 126 140

FT CARBOHYD 46 46

FT CARBOHYD 115 115

FT CARBOHYD 156 156

FT CARBOHYD 161 161

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200

FT CARBOHYD 200 200



```

RA MEDLINE; 93314973.
RX Triplido-Rosenthal H.G., Linser P.J., Greenberg R.M., Gleeson R.A.,
RA Carr W.E.;
RT "GONA clones from the olfactory organ of the spiny lobster encode a
RT protein related to eukaryotic glutamine synthetase.";
RL Gene 129:275-278(1993).
CC -1 CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) -> ADP + GLUTAMINE +
CC ORTHOPHOSPHATE.
CC -1 SUBUNIT: HOMODIMER.
CC -1 SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1 SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.sdb.ch).
CC
DR EMBL; M6798; AAA02583.1; -.
DR PIR; J00716; J00716.
DR HSSP; P06201; ZLGS.
DR PROSITE; PS00180; GLNA_1; 1.
DR PROSITE; PS00181; GLNA_ATP; 1.
DR PFAM; PF00120; gln-synt; 1.
DR KW
SQ SEQUENCE 361 AA; 40768 MW; 3D8C3C507676099C CRC64;
Query Match
Best Local Similarity 5.88; Score 92; DB 1; Length 361;
Matches 20; Conservative 22; Mismatches 26; Indels 6; Gaps 6;
Db 16 LDIPQCKAMVWV-DGTGENLRKSTRRLN-TPKSPSELPV-WNFGSGSTGAERNS 72
Qy 86 LD--AQRRDIIYFSIRNGHVPFTASRMWVYLKSPRHVFSWHODAVYVTCSDLSYG 144
Db 73 DYVLPAVNYRDPF 86
Qy 145 DL-LYEVO-YRSPF 156
RESULT 14
ID I131_HUMAN STANDARD; PRT; 427 AA.
AC P78552; O99656;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE INTERLEUKIN-13 RECEPTOR ALPHA-1 CHAIN PRECURSOR (IL-13R-ALPHA-1) (IL-
DE 13RA-1).
GN IL13RA1 OR IL13RA OR IL13R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CARCINOMA;
RX MEDLINE; 97165986.
RA Miloux B., Laurent P., Bonnin O., Lupker J., Caput D., Vita N.,
RA Ferreira P.;
RT "Cloning of the human IL-13R alpha chain and reconstitution with the
RT IL4R alpha of a functional IL-4/IL-13 receptor complex.";
RL FEBS Lett. 401:163-166(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=B-CELL;
RX MEDLINE; 97165986.
RA Gauchat J.F.M., Schlegenhaut E., Feng N.P., Moser R., Yamage M.,
RA Jeanin P., Alouani S., Elson G., Notarangelo L.D., Weiss T.,
RA Eugster H.P., Bonnefoy J.Y.;
RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=T-CELL;

```

```

RX MEDLINE: J9067184.
RA Aman M., Tayebi N., Obiri N.I., Puri R.K., Modi W.S.,
RA Leonard W.J.,
RT "CDNA cloning and characterization of the human interleukin 13
RT receptor alpha chain."
RL J. Biol. Chem. 271:29265-29270(1996).
CC -1- FUNCTION: BINDS IL-13 WITH A LOW AFFINITY. TOGETHER WITH IL-4R-
CC ALPHA CAN FORM A FUNCTIONAL RECEPTOR FOR IL-13. ALSO SERVES AS AN
CC ALTERNATE ACCESSORY PROTEIN TO THE COMMON CYTOKINE RECEPTOR GAMMA
CC CHAIN FOR IL-4 SIGNALING, BUT CANNOT REPLACE THE FUNCTION OF
CC GAMMA-C IN ALLOWING ENHANCED IL-2 BINDING ACTIVITY.
CC -1- SUBUNIT: INTERLEUKIN-13 RECEPTOR IS A COMPLEX OF IL4R-ALPHA,
CC IL13R-ALPHA, AND POSSIBLY OTHER COMPONENTS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: UBIQUITOUS. HIGHEST LEVELS IN HEART, LIVER,
CC SKELETAL MUSCLE AND OVARY; LOWEST LEVELS IN BRAIN, LUNG AND
CC KIDNEY. ALSO FOUND IN B-CELLS, T-CELLS AND ENDOTHELIAL CELLS.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.slb.ch/announce/
CC or send an email to license@sb.slb.ch).
CC -----
DR EMBL: Y10659; CAA71669.1; -
DR EMBL: Y09338; CAA70508.1; -
DR EMBL: U62858; AAB37127.1; -
DR HSSP: P31785; 11LN.
DR MIM: 300119; -
DR PFAM: PF00041; fn3; 1.
KW Receptor; Transmembrane; Glycoprotein; Immunoglobulin domain; Signal.
FT FT SIGNAL 1 21 POTENTIAL.
FT FT CHAIN 22 427 INTERLEUKIN-13 RECEPTOR ALPHA-1 CHAIN.
FT FT DOMAIN 22 343 EXTRACELLULAR (POTENTIAL).
FT FT TRANSMEM 344 367 CYTOPLASMIC (POTENTIAL).
FT FT DOMAIN 368 427 IG-LIKE C2-TYPE DOMAIN.
FT FT DISULFID 39 102 POTENTIAL.
FT FT DISULFID 46 95 BY SIMILARITY.
FT FT DISULFID 134 144 BY SIMILARITY.
FT FT DISULFID 173 185 BY SIMILARITY.
FT FT CARBOHYD 37 37 POTENTIAL.
FT FT CARBOHYD 61 61 POTENTIAL.
FT FT CARBOHYD 105 105 POTENTIAL.
FT FT CARBOHYD 138 138 POTENTIAL.
FT FT CARBOHYD 157 157 POTENTIAL.
FT FT CARBOHYD 235 235 POTENTIAL.
FT FT CARBOHYD 265 265 POTENTIAL.
FT FT CARBOHYD 293 293 POTENTIAL.
FT FT CARBOHYD 329 329 POTENTIAL.
FT FT CARBOHYD 341 341 POTENTIAL.
FT FT CONFLICT 130 130 T -> I (IN REF. 3).
FT FT CONFLICT 358 358 G -> D (IN REF. 3).
SQ SEQUENCE 1427 AA; 48759 MW; 5983B38BF554107B CRC64;
Query Match 5.7%; Score 90; DB 1; Length 427;
Best Local Similarity 30.9%; Pred. No. 5,93e-01;
Matches 21; Conservative 18; Mismatches 25; Indels 4; Gaps 4;
Db 122 EGDESAVTELOCIMWHLSTMKCSWLPGRNTSDPTNYLLYWHRSLEKIHOCEN-IFREG 180
QY 23 QGGAABEV-OIOLTFVETLEWVOTWASRY-SR-TNLFHFYRNGDEAVDQCTNVLQDG 79
Db 181 QYEGCSPD 188
QY 80 HTSGCLD 87

```



```

ID KS6B_XENLA STANDARD; PRT; 629 AA.
AC P10666;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE RIBOSOMAL PROTEIN S6 KINASE II BETA (EC 2.7.1.-) (S6KII-BETA)
DE (P90-RSK).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88217904.
RA Jones S.W., Erikson E., Blenis J., Maller J.L., Erikson R.L.;
RT "A Xenopus ribosomal protein S6 kinase has two apparent kinase
RL domains that are each similar to distinct protein kinases.";
Proc. Natl. Acad. Sci. U.S.A. 85:3377-3381(1988).
CC -!- FUNCTION: PHOSPHORYLATES A WIDE RANGE OF SUBSTRATES INCLUDING
CC RIBOSOMAL PROTEIN S6. IMPLICATED IN THE ACTIVATION OF THE MITOGEN-
CC ACTIVATED KINASE CASCADE.
CC -!- SIMILARITY: BELONGS TO THE PROTEIN KINASE SUPERFAMILY; SER/THR
CC FAMILY. CONTAINS TWO KINASE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M20188; AAA49959.1;
DR PIR: A30001; A30001.
DR HSP: O63450; IA06.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 2.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 2.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 2.
DR PFAM: PF00069; pkinase; 2.
DR PFAM: PF00433; pkinase_C; 1.
DR Transferrase: Serine/threonine-protein kinase; ATP-binding;
KW Repeat; Multigene family.
FT DOMAIN 62 321
FT DOMAIN 416 629
FT NP_BIND 68 76
FT BINDING 94 94
FT ACT_SITE 187 187
FT NP_BIND 422 430
FT BINDING 445 445
FT ACT_SITE 533 533
FT ACT_SITE 533 533
FT ACT_SITE 533 533
SQ SEQUENCE 629 AA; 71286 MW; BC6144415F1DCE0 CRC64;
PROTEIN KINASE 1.
PROTEIN KINASE 2.
ATP (BY SIMILARITY).
ATP (BY SIMILARITY).
BY SIMILARITY.
ATP (BY SIMILARITY).
ATP (BY SIMILARITY).
BY SIMILARITY.

```

Query Match 5.78; Score 90; DB 1; Length 629;

Best Local Similarity 31.78; Pred. No. 5.93e-01;

Matches 19; Conservative 14; Mismatches 23; Indels 4; Gaps 3;

Db 475 DYTEENSTYLVTELMRGGLRLRQK--FFSERACSVLFTVCK-TVEYLSQGVVH 531

QY 66 EAYDQCTN-YLQEGHTSGCLLDAEQRDILYFSIRNGTHEVFTASRMVYILKPSPKH 124

Search completed: Wed May 10 11:30:18 2000

Job time : 109 secs.



\*\*\*\*\*  
 M P S R C H  
 \*\*\*\*\*  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 11:30:35 2000; Maspar time 240.95 Seconds  
 60.139 Million cell updates/sec

Tabular output not generated.

Title: >US-09-376-430-2

Description: (23-231) from US09376430A.pep (4 of 25)

Perfect Score: 1586

Sequence: 1 QGGAEGVQIIYFNLETV.....GEIRDACAEPTTPPKLKS 209

Scoring table: PAM 150

Gap 11

Searched: 225878 segs, 69334122 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: sptrembl12

1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human

5:sp.invertebrate 6:sp.mammal 7:sp.mhc 8:sp.organelle

9:sp.phage 10:sp.plant 11:sp.rodent 12:sp.unclassified

13:sp.vertbrate 14:sp.virus

Statistics: Mean 42.952; Variance 69.764; scale 0.616

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	130	8.2	383	11	IL-13 RECEPTOR ALPHA 2	2.33e-07
2	109	6.9	435	2	PUTATIVE ASPARTATE AMI	1.19e-03
3	109	6.9	435	2	ASPARTATE AMINOTRANSFE	1.19e-03
4	100	6.3	890	11	INTERLEUKIN-5 RECEPTOR	3.56e-02
5	99	6.2	4436	1	4436AA LONG HYPOTHETIC	5.13e-02
6	99	6.2	6048	5	UNC-22 PROTEIN.	5.13e-02
7	99	6.2	6831	5	UNC-22 PROTEIN.	5.13e-02
8	99	6.2	7160	5	UNC-22 PROTEIN.	5.13e-02
9	97	6.1	256	11	CYTOKINE RECEPTOR COMP	1.06e-01
10	97	6.1	311	2	SIMILAR TO TRIMETHYLAM	1.06e-01
11	96	6.1	936	11	INTERLEUKIN-3 RECEPTOR	1.52e-01
12	95	6.0	315	11	MOR 5/BETA3	2.17e-01
13	95	6.0	638	6	GROWTH HORMONE RECEPTO	2.17e-01
14	95	6.0	673	14	POTATIVE RNA DEPENDENT	2.17e-01
15	95	6.0	987	5	SIMILAR TO IMMUNOGLOBU	2.17e-01
16	93	5.9	173	5	COSMID C29F5.	4.39e-01
17	94	5.9	407	5	PROBABLE EPH-LIKE KINA	3.09e-01
18	94	5.9	919	5	POSSIBLE DEHYDROGENAS	3.09e-01
19	94	5.9	981	3	EPH RECEPTOR TYROSINE	3.09e-01
20	94	5.9	1122	5	EPH RECEPTOR TYROSINE	3.09e-01

21	94	5.9	26926	4	Q10466	TITIN, HEART ISOFORM N	3.09e-01
22	92	5.8	202	14	P90279	NEF PROTEIN.	6.22e-01
23	92	5.8	684	13	Q90544	NOVEL ANTIGEN RECEPTOR	6.22e-01
24	91	5.7	1395	5	O44924	ROUNDABOUT 1.	8.78e-01
25	92	5.8	202	14	P88435	NEF PROTEIN.	1.24e+00
26	90	5.7	204	14	O74919	NEF.	1.24e+00
27	90	5.7	204	14	O74920	NEF.	1.24e+00
28	90	5.7	204	14	O74921	NEF.	1.24e+00
29	90	5.7	256	10	O65756	VEGETATIVE LECTIN.	1.24e+00
30	90	5.7	344	5	P91017	COSMID C01G8.	1.24e+00
31	90	5.7	375	1	O95896	375AA LONG HYPOTHETICA	1.24e+00
32	90	5.7	427	4	O95646	INTERLEUKIN-13 RECEPTO	1.24e+00
33	90	5.7	935	2	O31000	INTIMIN.	1.24e+00
34	89	5.6	935	2	O3X207	CONSERVED HYPOTHETICAL	1.74e+00
35	89	5.6	339	2	O92199	POTATIVE HISTIDINE KIN	1.74e+00
36	89	5.6	634	6	O46600	SOMATOTROPIN RECEPTOR	1.74e+00
37	89	5.6	1286	10	O92R72	P-GLYCOPROTEIN.	1.74e+00
38	89	5.6	1466	3	O42930	VACUOLAR PROTEIN SORTI	1.74e+00
39	88	5.5	137	14	Q07344	STRUCTURAL PROTEIN P11	2.44e+00
40	88	5.5	269	6	P79195	GROWTH HORMONE BINDING	2.44e+00
41	88	5.5	366	3	O42858	HYPOTHETICAL 41.7 KD T	2.44e+00
42	88	5.5	459	2	O86960	ALPHA-GLUCOSIDASE.	2.44e+00
43	87	5.5	563	14	Q9WT04	GANCICLOVIR KINASE, CO	3.40e+00
44	88	5.5	934	2	O85627	L0025.	2.44e+00
45	87	5.5	938	2	O52620	INTIMIN.	3.40e+00

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	383 AA.
ID	O88786			
AC	O88786;1998	(Tremblrel. 08, Created)		
DT	01-NOV-1998	(Tremblrel. 08, Last sequence update)		
DR	01-NOV-1998	(Tremblrel. 08, Last annotation update)		
DE	IL-13 RECEPTOR ALPHA 2.			
OS	Mus musculus (Mouse)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OT	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C3H/HEJ; TISSUE=THYMUS;			
RX	MEDLINE; 98391042.			
RA	DONALDSON D.D., WHITTERS M.J., FITZ L., NEBEN T.Y., FINNERTY H.,			
RA	HENDERSON S.L., O'HARA R.M. JR., BEIER D.R., TURNER K.J., WOOD C.R.,			
RA	COLLINS M.			
RT	"The murine IL-13 receptor alpha 2: molecular cloning,			
RT	characterization, and comparison with murine IL-13 receptor alpha 1."			
EL	J. Immunol. 161:2317-2324(1998).			
DR	EMBL; U65747; AAC33240.1;			
SQ	SEQUENCE 383 AA; 44483 MW; 5EAEP3E3 CRC32;			

Query Match	8.2%;	Score 130;	DB 11;	Length 383;
Best Local Similarity	30.2%;	Pred. No. 2.33e-07;		
Matches	26;	Conservative	22;	Mismatches 32; Indels 6; Gaps 5;
Db	127	EGSLETKIQMKCIYNNQYLVCSWKPGKTVISDTNFMFWYGLDHAL-OCADYLQHD 185		
Qy	23	QCGAAGVQ-IQIIYFNLETVQVITWASK--YSRTNLFTHYRFG-DEAYDQCTNYLQE 78		
Db	186	EKNVGCKLSNLDSSDYKDFFCVNGS 211		
Qy	79	GHTSGCLLDAEQRDDIL-YFSIRNGT 103		

RESULT	2	PRELIMINARY;	PRT;	435 AA.
ID	O06685			
AC	O06685			
DT	01-JUL-1997	(Tremblrel. 04, Created)		
DR	01-JUL-1997	(Tremblrel. 04, Last sequence update)		
DE	01-NOV-1998	(Tremblrel. 08, Last annotation update)		
DE	POTATIVE ASPARTATE AMINOTRANSFERASE TPAAT.			
GN	TPAAT.			

```

OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-PALLIDUM.
RX MEDLINE; 97294457.
RA SHEVCHENKO D.V., AKINS D.R., ROBINSON E., LI M., POPOVA T.G.,
RA COX D.L., RADOLF J.;
RT "Molecular characterization and cellular localization of TplRR, a
RT processed leucine-rich repeat protein of Treponema pallidum, the
RT syphilis spirochete.";
RL J. Bacteriol. 179:3188-3195(1997).
DR EMBL; U073748; AAC45303.1; -.
SQ Transferase; Aminotransferase.
SQ SEQUENCE 435 AA; 47285 MW; 7AF01306 CRC32;

Query Match 6.9%; Score 109; DB 2; Length 435;
Best Local Similarity 35.1%; Pred.No.1.19e-03;
Matches 20; Conservative 11; Mismatches 21; Indels 5; Gaps 5;

Db 206 AVLVICDD-AYSGFEYEASLMRGSFFARFAQAHNKICALKIDGLTKEE-YA-WGLRV 259
QY 133 AVTVTCSDLISYGDLLYEVOY-RSPFTEWQSKQENTCNVTIEGLDAEKCYSEW-VRV 187

RESULT 3
ID O83252 PRELIMINARY; PRT; 435 AA.
AC O83252;
RX MEDLINE; 98332770.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.,
RA DODSON R., GWNN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.,
RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTERBACK T.,
RA MCDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.,
RA VENTER J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
[2]
RN SEQUENCE FROM N.A.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.,
RA DODSON R., GWNN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.,
RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTERBACK T.,
RA MCDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.,
RA VENTER J.C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE001204; AAC65212.1; -.
DR TIGR; TP0223; -.
SQ Transferase; Aminotransferase.
SQ SEQUENCE 435 AA; 47304 MW; 0B31A6E1 CRC32;

Query Match 6.9%; Score 109; DB 2; Length 435;
Best Local Similarity 35.1%; Pred.No.1.19e-03;
Matches 20; Conservative 11; Mismatches 21; Indels 5; Gaps 5;

Db 206 AVLVICDD-AYSGFEYEASLMRGSFFARFAQAHNKICALKIDGLTKEE-YA-WGLRV 259
QY 133 AVTVTCSDLISYGDLLYEVOY-RSPFTEWQSKQENTCNVTIEGLDAEKCYSEW-VRV 187

RESULT 4

```



```

DE CYTOKINE RECEPTOR COMPLEX COMMON BETA CHAIN H BETA C (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94235843.
RA D'ANDREA R., RAYNER J., MORETTI P., LOPEZ A., GOODALL G.J.,
RA GONDA T.J., VADAS M.;
RT "A mutation of the common receptor subunit for interleukin-3 (IL-3),
RT granulocyte-macrophage colony-stimulating factor, and IL-5 that leads
RT to ligand independence and tumorigenicity."
RL Blood 83:2802-2808(1994).
DR EMBL; S70302; AAB31055.1; -
DR HSSP; P40189; 1BQU.
DR PFAM; PF00041; fn3; 1.
FT NON_TER
SQ SEQUENCE 256 AA; 29206 MW; 5242B76B CRC32;

Query Match 6.1%; Score 97; DB 11; Length 256;
Best Local Similarity 22.6%; Pred. No. 1.06e-01;
Matches 14; Conservative 16; Mismatches 29; Indels 3; Gaps 3;

Db 80 RTGNGINSESEAH-SMALPALEPSTRIWVRVTSRTGYN-GIW-SEWSEARSNDTE 136
QY 153 RSPDTEWQSQENTCNVTIEGLDAEKCYSEFWVRVKAMEDVYGPDPSPDSEVTCWORG 212
Db 137 SV 138
QY 213 EI 214

RESULT 10
ID Q48962 PRELIMINARY; PRT; 311 AA.
AC Q48962;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE SIMILAR TO TRIMETHYLAMINE DH (FRAGMENT).
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC capricolum group.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 27343(KID);
RX MEDLINE; 96059641.
RA BORK P., OUZOUNIS C., CASARI G., SCHNEIDER R., SANDER C., DOLAN M.,
RA GILBERT W., GILLET P.M.;
RT "Exploring the Mycoplasma capricolum genome: a minimal cell reveals
RT its physiology."
RL Mol. Microbiol. 16:955-967(1995).
DR EMBL; Z33015; CAA83700.1; -
DR PFAM; PF00724; oxidored_FMN; 1.
FT NON_TER 311
SQ SEQUENCE 311 AA; 35768 MW; 6CC72B66 CRC32;

Query Match 6.1%; Score 97; DB 2; Length 311;
Best Local Similarity 27.9%; Pred. No. 1.06e-01;
Matches 12; Conservative 11; Mismatches 19; Indels 1; Gaps 1;

Db 202 FCELVVKALREVIDKYAPKNFIFGRATPEEYGDILGYTIED 244
QY 37 FNLVQVTVWNA-SKYSRTNLTFHFRNGDEAYDOCTNYLLOE 78

RESULT 11
ID Q64146 PRELIMINARY; PRT; 896 AA.
AC Q64146;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE INTERLEUKIN-3 RECEPTOR BETA-SUBUNIT (FRAGMENT).
GN RIL-3<BETA>

```

```

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95370942.
RA APPEL K., BUTTINI M., SAUTER A., GEBICKE-HARTER P.J.;
RT "Cloning of rat interleukin-3 receptor beta-subunit from cultured
RT microglia and its mRNA expression in vivo."
RL J. Neurosci. 15:5800-5809(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-WISTAR; TISSUE-BRAIN;
RA GEBICKE-HARTER P.J.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; S79263; AAB35068.1; -
DR EMBL; AJ000555; CAA04186.1; -
DR PFAM; PF00041; fn3; 2.
KW Signal.
FT NON_TER 896
FT NON_TER 896
SQ SEQUENCE 896 AA; 99504 MW; 8E7ED2CD CRC32;

Query Match 6.1%; Score 96; DB 11; Length 896;
Best Local Similarity 30.3%; Pred. No. 1.52e-01;
Matches 20; Conservative 17; Mismatches 21; Indels 8; Gaps 8;

Db 375 FOYQKKKLDLDR-WEDSKTENLNHAHMDLPOLPEGTSCARVRVKTYPE-Y-KGLW-SEW 430
QY 148 YEVOYRSPFDTEWQ-SKOENTCNV-TIE-G-LDAEKCYSEFWVRVKAMEDVYGPDPSPDW 203
Db 431 SNECTW 436
QY 204 SEVTCW 209

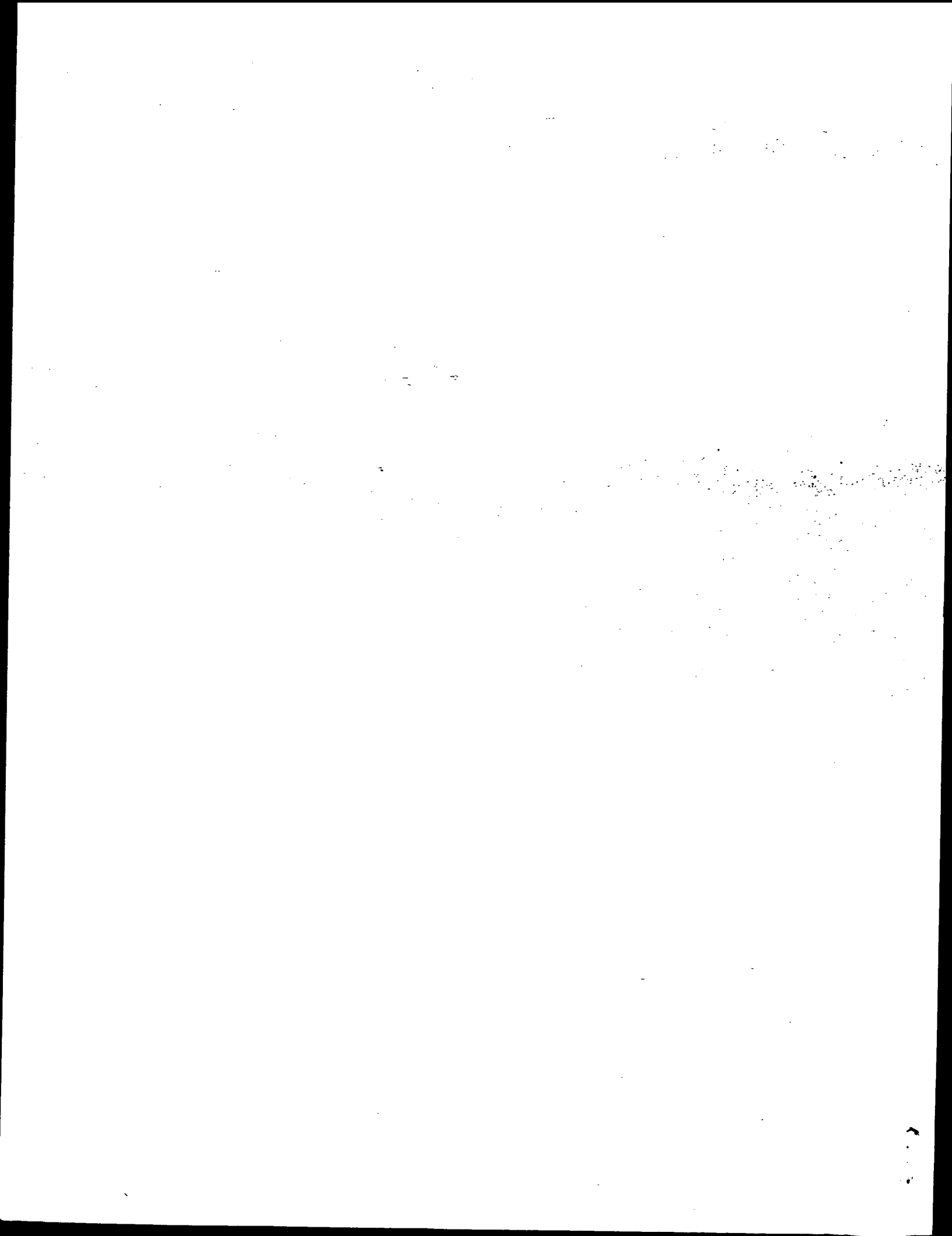
RESULT 12
ID Q9WVN6 PRELIMINARY; PRT; 315 AA.
AC Q9WVN6;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE MOR 5'BETA3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-129;
RX MEDLINE; 99055560.
RA BENDER M.A., REIK A., CLOSE J., TELLING A., EPNER E., FIERING S.,
RA HARDISON R., GROUDINE M.;
RT "Description and targeted deletion of 5' hypersensitive site 5 and 6
RT of the mouse beta-globin locus control region."
RL Blood 92:4394-4403(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-129;
RX MEDLINE; 99238494.
RA BULGER M., VON DOORNICK J.H., SAITOH N., TELLING A., FARRELL C.,
RA BENDER M.A., FELSENFELD G., AXEL R., GROUDINE M.;
RT "Conservation of sequence and structure flanking the mouse and human
RT beta-globin loci: the beta-globin genes are embedded within an array
RT of odorant receptor genes."
RL Proc. Natl. Acad. Sci. U.S.A. 96:5129-5134(1999).
DR EMBL; AF074080; AAD28304.1; -
SQ SEQUENCE 315 AA; 35559 MW; DEE53BF7 CRC32;

Query Match 6.0%; Score 95; DB 11; Length 315;
Best Local Similarity 28.3%; Pred. No. 2.17e-01;
Matches 13; Conservative 13; Mismatches 17; Indels 3; Gaps 3;

Db 159 PIIRLHWFPYC-RSHVLSHA-FCLHQDVIKLACADITF-NRLYPV 201

```











US-09-376-430-2-05.rag

Thu May 11 06:49:34 2000

(SUGA/) SUGAMURA K.  
 PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
 PI Suzuki M, Takeshita T;  
 DR WPI: 94-017546/03.  
 DR P-PSDB: Q54830.  
 DNA and protein sequences of IL-2 gamma chain - useful as immune  
 PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
 PT transplant rejection  
 PS Disclosure: Page 21-22, 34-35; 50pp; English.  
 CC The human IL-2 receptor gamma chain preform (R47148), including the  
 CC signal peptide, is encoded by the sequence given in Q54828. The  
 CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
 CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830.  
 CC while a soluble form suitable for expression in prokaryotes (R47151)  
 CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
 CC terminal sequence of IL-2 receptor gamma chain, and are used to  
 CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
 CC are used to obtain the protein given in R47151.  
 SQ Sequence 252 AA;  
 Query Match 12.6%; Score 195; DB 1; Length 252;  
 Best Local Similarity 25.9%; Pred. No. 6.64e-09;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 59 EVQCFFVFNVEYMNCTWNSSEPOPTNLTHLYWYKNSDNDKVKCSHYLSEETISGCOLQ 118  
 QY 31 QIQIIFNLETVQVWNASYSR-TNLTFHYRF-NGD-EAYDQCTNLLQEGHTSGCLLD 87  
 Db 119 KKEHLVQTFVQLODPRPRQATQMLKQNLVWPAPENLTLLKLSOLELNNWNR 178  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRWV-VYIL-KPSSPKHVRP-SWHQDAVTVTCSDL 142  
 Db 179 LNHCLHVLQVYRTDMDHSHWTEQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 237  
 QY 143 YGD-LLEYQVRSFDTW-OSKQENTCNVTIEGLDAEKCSFWVRKAMEDVYGPDTYP 200  
 Db 238 SEWSHPHWH 246  
 QY 201 SDWSEVTCW 209  
 RESULT 5  
 ID R47149 standard; Protein; 347 AA.  
 AC R47149;  
 DT 13-JUN-1994 (first entry)  
 DE IL-2 receptor gamma chain.  
 KW Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;  
 KW rheumatoid arthritis; transplant rejection; primer;  
 KW polymerase chain reaction; PCR; amplification.  
 OS Homo sapiens.  
 PN Ep-578932-A.  
 PD 19-JAN-1994.  
 PE 23-APR-1993; JP-104947.  
 PR 23-APR-1992; JP-104947.  
 PA (AJIN) AJINOMOTO KK.  
 PA (SUGA/) SUGAMURA K.  
 PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
 PI Suzuki M, Takeshita T;  
 DR WPI: 94-017546/03.  
 DR P-PSDB: Q54829.  
 DNA and protein sequences of IL-2 gamma chain - useful as immune  
 PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
 PT transplant rejection  
 PS Claim 4; Page 41; 50pp; English.  
 CC The human IL-2 receptor gamma chain preform (R47148), including the  
 CC signal peptide, is encoded by the sequence given in Q54828. The  
 CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
 CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830.  
 CC while a soluble form suitable for expression in prokaryotes (R47151)  
 CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
 CC terminal sequence of IL-2 receptor gamma chain, and are used to  
 CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
 CC are used to obtain the protein given in R47151.

SQ Sequence 347 AA;  
 Query Match 12.6%; Score 195; DB 1; Length 347;  
 Best Local Similarity 25.9%; Pred. No. 6.64e-09;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 37 EVQCFFVFNVEYMNCTWNSSEPOPTNLTHLYWYKNSDNDKVKCSHYLSEETISGCOLQ 96  
 QY 31 QIQIIFNLETVQVWNASYSR-TNLTFHYRF-NGD-EAYDQCTNLLQEGHTSGCLLD 87  
 Db 97 KKEHLVQTFVQLODPRPRQATQMLKQNLVWPAPENLTLLKLSOLELNNWNR 156  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRWV-VYIL-KPSSPKHVRP-SWHQDAVTVTCSDL 142  
 Db 157 LNHCLHVLQVYRTDMDHSHWTEQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 215  
 QY 143 YGD-LLEYQVRSFDTW-OSKQENTCNVTIEGLDAEKCSFWVRKAMEDVYGPDTYP 200  
 Db 216 SEWSHPHWH 224  
 QY 201 SDWSEVTCW 209  
 RESULT 6  
 ID R47148 standard; Protein; 369 AA.  
 AC R47148;  
 DT 13-JUN-1994 (first entry)  
 DE IL-2 receptor gamma chain.  
 KW Interleukin-2 receptor gamma chain; IL-2; receptor; immunoregulator;  
 KW rheumatoid arthritis; transplant rejection; primer;  
 KW polymerase chain reaction; PCR; amplification; ss.  
 OS Homo sapiens.  
 PN Ep-578932-A.  
 PD 19-JAN-1994.  
 PE 22-APR-1993; 106561.  
 PR 23-APR-1992; JP-104947.  
 PA (AJIN) AJINOMOTO KK.  
 PA (SUGA/) SUGAMURA K.  
 PI Asao H, Hamuro J, Nakamura M, Shimamura T, Sugamura K;  
 PI Suzuki M, Takeshita T;  
 DR WPI: 94-017546/03.  
 DR P-PSDB: Q54828.  
 DNA and protein sequences of IL-2 gamma chain - useful as immune  
 PT regulatory agents for treatment of e.g. rheumatoid arthritis and  
 PT transplant rejection  
 PS Disclosure: Page 16-17, 29-30; 50pp; English.  
 CC The human IL-2 receptor gamma chain preform (R47148), including the  
 CC signal peptide, is encoded by the sequence given in Q54828. The  
 CC mature protein (R47149) is encoded by sequence Q54829. A soluble  
 CC form of IL-2 receptor gamma chain (R47150) is encoded by Q54830.  
 CC while a soluble form suitable for expression in prokaryotes (R47151)  
 CC is encoded by Q54831. Primers 1-6 (Q54820-25) are based on the N-  
 CC terminal sequence of IL-2 receptor gamma chain, and are used to  
 CC isolate IL2 receptor gamma chain receptor cDNA. Primers Q54826-27  
 CC are used to obtain the protein given in R47151.  
 SQ Sequence 369 AA;  
 Query Match 12.6%; Score 195; DB 1; Length 369;  
 Best Local Similarity 25.9%; Pred. No. 6.64e-09;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 59 EVQCFFVFNVEYMNCTWNSSEPOPTNLTHLYWYKNSDNDKVKCSHYLSEETISGCOLQ 118  
 QY 31 QIQIIFNLETVQVWNASYSR-TNLTFHYRF-NGD-EAYDQCTNLLQEGHTSGCLLD 87  
 Db 119 KKEHLVQTFVQLODPRPRQATQMLKQNLVWPAPENLTLLKLSOLELNNWNR 178  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRWV-VYIL-KPSSPKHVRP-SWHQDAVTVTCSDL 142  
 Db 179 LNHCLHVLQVYRTDMDHSHWTEQSDYRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 237

Expression vector; pME18S/mFas.EXT-AIC2A; murine Fas antigen;

KW extracellular; region; AIC2A; soluble membrane protein; antibody production; diseases; treatment; prevention.

OS Mus musculus.

PN J07115988-A.

PD 09-MAY-1995.

PF 26-OCT-1993; 267644.

PR 26-OCT-1993; JP-267644.

PA (NIBS ) JAPAN TOBACCO INC.

DR WPI; 95-202847/27.

DR N-PSDB; 095306.

PT Preparation of soluble membrane proteins - for their use in antibody production for the treatment and prevention of related diseases

PS Claim 10; pages 36-38; 51pp; Japanese.

CC R78616 is the protein prod. of the expression vector pME18S/murine Fas antigen; extracellular region-AIC2A. The expression vector was used for the prodn. of recombinant soluble membrane proteins. The proteins can be used in antibody prodn. for the treatment and prevention of related diseases.

CC diseases.

SQ Sequence 576 AA;

Query Match 7.4%; Score 115; DB 1; Length 576;  
Best Local Similarity 32.4%; Pred. No. 1.64e-01;  
Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 502 YIDHTFQVQYKKKSES-WKDSKTENLGRVNSMDLPQLEPDTSYCARVRVKPISD-YD-GI 558  
QY 143 YGDLLEYQVRSFDTETW-OSKQENTCNV-TIE-G-LDAEKCYSFVVRKAMEDVYGPDT 198

Db 559 W-SEWSNEYTW 568  
QY 199 YPSDWSEVTCW 209

RESULT 10

ID R78616 standard; Protein; 596 AA.

AC R78616;

DT 20-FEB-1996 (first entry)

DE Expression vector pME18S/MPD-1.EXT-AIC2A protein prod.

PS Claim 10; Pages 44-46; 51pp; Japanese.

KW Expression vector; pME18S/MPD-1.EXT-AIC2A; soluble membrane protein; extracellular; region; AIC2A; soluble membrane protein; antibody production; diseases; treatment; prevention.

KW Mus musculus.

OS Mus musculus.

PN J07115988-A.

PD 09-MAY-1995.

PF 26-OCT-1993; 267644.

PR 26-OCT-1993; JP-267644.

PA (NIBS ) JAPAN TOBACCO INC.

DR WPI; 95-202847/27.

DR N-PSDB; 095310.

PT Preparation of soluble membrane proteins - for their use in antibody production for the treatment and prevention of related diseases

PS Claim 10; Pages 44-46; 51pp; Japanese.

CC R78616 is the protein prod. of the expression vector pME18S/MPD-1 .extracellular region-AIC2A. The expression vector was used for the prodn. of recombinant soluble membrane proteins. The proteins can be used in antibody prodn. for the treatment and prevention of related diseases.

CC diseases.

SQ Sequence 596 AA;

Query Match 7.4%; Score 115; DB 1; Length 596;  
Best Local Similarity 32.4%; Pred. No. 1.64e-01;  
Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 517 YIDHTFQVQYKKKSES-WKDSKTENLGRVNSMDLPQLEPDTSYCARVRVKPISD-YD-GI 573  
QY 143 YGDLLEYQVRSFDTETW-OSKQENTCNV-TIE-G-LDAEKCYSFVVRKAMEDVYGPDT 198

Db 574 W-SEWSNEYTW 583  
QY 199 YPSDWSEVTCW 209

RESULT 11

R92526 standard; Protein; 600 AA.

R92526;

AC R92526 (first entry)

DE Fas antigen #1.

OS Fas antigen; immunoassay; monoclonal antibody; autoimmune disease; SLE; rheumatoid arthritis; serum; systemic lupus erythematosus.

KW Synthetic.

OS Synthetic.

PH Key

FT 1.16 Location/Qualifiers

FT peptide

FT protein

FT 17.600 /note= "signal peptide"

FT 17.600 /note= "mature Fas antigen #1"

PN W09601277-A1.

PD 18-JAN-1996. J00349.

PF 03-MAR-1995; JP-154706.

PR 06-JUL-1994; JP-025637.

PR 14-FEB-1995; JP-025637.

PA (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.

PA (NIBS ) JAPAN TOBACCO INC.

PI Hachiya T, Noguchi J, Yonehara S;

DR WPI; 96-087635/09.

DR N-PSDB; T16300.

PT Immunoassay method for soluble Fas antigen in body fluids - for diagnosis of autoimmune diseases such as rheumatoid arthritis and systemic lupus erythematosus

PS Claim 13; Page 73-77; 124pp; Japanese.

CC R92526 and R92527 represent soluble Fas antigens. These soluble Fas antigen is included in the immunoassay kit of the invention. The kit is for the assay of soluble Fas antigen and contains an immobilised anti-soluble Fas monoclonal antibody, as well as one of these standard soluble Fas antigens. The assay is simple and has high accuracy, high sensitivity, and is capable of assaying a number of different specimens at the same time. The immunoassay is used on biological samples (such as serum) and is useful for diagnosis of autoimmune diseases such as rheumatoid arthritis or systemic lupus erythematosus (SLE).

CC Sequence 600 AA;

Query Match 7.4%; Score 115; DB 1; Length 600;  
Best Local Similarity 32.4%; Pred. No. 1.64e-01;  
Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 521 YIDHTFQVQYKKKSES-WKDSKTENLGRVNSMDLPQLEPDTSYCARVRVKPISD-YD-GI 577  
QY 143 YGDLLEYQVRSFDTETW-OSKQENTCNV-TIE-G-LDAEKCYSFVVRKAMEDVYGPDT 198

Db 578 W-SEWSNEYTW 587  
QY 199 YPSDWSEVTCW 209

RESULT 12

ID R78610 standard; Protein; 600 AA.

AC R78610;

DT 19-FEB-1996 (first entry)

DE Expression vector pME18S/hFas.EXT-AIC2A protein prod.

PS Claim 10; Pages 44-46; 51pp; Japanese.

KW Expression vector; pME18S/hFas.EXT-AIC2A; human Fas antigen; extracellular; region; AIC2A; soluble membrane protein; antibody production; diseases; treatment; prevention.

KW Homo sapiens.

OS Homo sapiens.

PH Key

FT 1.16 Location/Qualifiers

FT peptide

FT 17.600 /label= sig\_peptide

FT peptide

FT 17.600 /label= mat\_peptide

PN J07115988-A.

PD 09-MAY-1995.

PF 26-OCT-1993; 267644.

PR 26-OCT-1993; JP-267644.

PA (NIBS ) JAPAN TOBACCO INC.

DR WPI; 95-202847/27.

DR N-PSDB; 095301.

PT Preparation of soluble membrane proteins - for their use in antibody production for the treatment and prevention of related diseases

PS Claim 10; Pages 28-30; 51pp; Japanese.  
 CC R78610 is the protein prod. of the expression vector pME18S/human Ras  
 CC antigen-extracellular region-AIC2A. The expression vector was used for  
 CC the prodn. of recombinant soluble membrane proteins. The proteins can  
 CC be used in antibody prodn. for the treatment and prevention of related  
 CC diseases.  
 SQ Sequence 600 AA;

Query Match 7.4%; Score 115; DB 1; Length 600;  
 Best Local Similarity 32.4%; Pred. No. 1.64e-01;  
 Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 521 YIDHTFOVQYKKKSS-WKDSKTENLGRVNSMDLPQLEPDTSYCARVRKPISD-YD-GI 577  
 QY 143 YGDLLEYQVRSFDTFTEW-QSKQENTCNV-TIE-G-LDAEKCYSFWRVKAMEDVYGPDT 198  
 Db 578 W-SEWSNEYTW 587  
 QY 199 YPSDWSNTEVTCW 209

RESULT 13  
 ID R78608 standard; Protein; 878 AA.

AC R78608;  
 DT 19-FEB-1996 (first entry)  
 DE Murine IL-3 receptor beta-subunit fragment AIC2A.  
 KW Plasmid pAIC2-26; murine AIC2A cDNA; soluble membrane protein;  
 KW antibody production; diseases; treatment; prevention;  
 KW interleukin-3; IL-3; receptor; beta-subunit; fragment.  
 OS Mus musculus.  
 FH Key Location/Qualifiers  
 FT peptide 1. 22  
 FT peptide /label= sig\_peptide  
 FT peptide 23. 878  
 FT peptide /label= mat\_peptide  
 PN J07115988-A.  
 PD 09-MAY-1995.  
 PF 26-OCT-1993; 267644.  
 PR 26-OCT-1993; JP-267644.  
 PA (NIBS) JAPAN TOBACCO INC.  
 DR WPI; 95-202847/27.  
 DR N-PSDB; Q95299.  
 DT Preparation of soluble membrane proteins - for their use in antibody  
 PT production for the treatment and prevention of related diseases  
 PS Example 1; Pages 19-22; 51pp; Japanese.  
 CC R78608 (the murine IL-3 receptor beta-subunit fragment AIC2A) is  
 CC encoded by the plasmid pAIC2-26, which contains the AIC2A cDNA. The  
 CC plasmid was used in the construction of an expression vector for the  
 CC prodn. of recombinant soluble membrane proteins. The proteins can be  
 CC used in antibody prodn. for the treatment and prevention of related  
 CC diseases.  
 SQ Sequence 878 AA;

Query Match 7.4%; Score 115; DB 1; Length 878;  
 Best Local Similarity 32.4%; Pred. No. 1.64e-01;  
 Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 370 YIDHTFOVQYKKKSS-WKDSKTENLGRVNSMDLPQLEPDTSYCARVRKPISD-YD-GI 426  
 QY 143 YGDLLEYQVRSFDTFTEW-QSKQENTCNV-TIE-G-LDAEKCYSFWRVKAMEDVYGPDT 198  
 Db 427 W-SEWSNEYTW 436  
 QY 199 YPSDWSNTEVTCW 209

RESULT 14  
 ID R20982 standard; Protein; 897 AA.

AC R20982;  
 DT 17-MAY-1992 (first entry)  
 DE Sequence of beta-chain of a human granulocyte-macrophage  
 DE colony stimulating factor (GM-CSF) receptor.  
 KW Agonist; antagonist; myeloid leukaemia; therapy; screening;

KW diagnosis; granulocyte-macrophage colony stimulating factor.  
 OS Homo sapiens.

FH Key Location/Qualifiers  
 FT peptide 1. 17  
 FT modified\_site /label= signal  
 FT modified\_site 58. 60  
 FT modified\_site /label= glycosylation  
 FT modified\_site 191. 193  
 FT modified\_site /label= as above  
 FT modified\_site 346. 348  
 FT modified\_site /label= as above

PN WO9201788-A.  
 PD 06-FEB-1992.  
 PE 16-JUL-1991; U04846.  
 PR 18-JUL-1990; US-554745.  
 PA (SCHE) SCHERING CORP.  
 PI Hayashida K, Kitamura T, Miyajima A;  
 DR WPI; 92-064947/08.  
 DR N-PSDB; Q21453.

PT Beta-chain of human granulocyte-macrophage CSF receptor - used  
 PT for screening agonists and antagonists of human GM-CSF, e.g. for  
 PT diagnosing myeloid leukaemia  
 PS Claim 2; Page 15-18; 26pp; English.  
 CC The nucleic acid encoding the human GM-CSF beta-chain is isolated  
 CC from a cDNA library prepd. from poly(A)+ RNA from TF-1 cells. The  
 CC high affinity human GM-CSF receptor (Kd < 1 nM) can be used for  
 CC screening candidate GM-CSF agonists and antagonists e.g. for  
 CC treating myeloid leukaemias.  
 SQ Sequence 897 AA;

Query Match 7.3%; Score 113; DB 1; Length 897;  
 Best Local Similarity 31.9%; Pred. No. 2.43e-01;  
 Matches 22; Conservative 10; Mismatches 34; Indels 3; Gaps 3;

Db 372 FEIQYRKDTAT-WKDSKTETLQNAHSMALPALEPSTFYARVRKRTSGYNGIWSEWSE 430  
 QY 148 YEVOYRSPFDTFTEW-QSKQENTCNVTIEGLDA-EKCYSFWRVKAMEDVYGPDTVPDSE 205  
 Db 431 ARSWDTEW 439  
 QY 206 VTCWQGEI 214

RESULT 15

ID R22229 standard; Protein; 237 AA.  
 AC R22229;  
 DT 20-JUL-1992 (first entry)  
 DE Truncated human growth hormone binding protein.  
 KW Soluble hGHbp; zinc finger; chelate; hGH receptor;  
 KW receptor-ligand complex.  
 OS Homo sapiens.  
 PN WO9203478-A.  
 PD 05-MAR-1992.  
 PF 16-AUG-1991; U05856.  
 PR 17-AUG-1990; US-568936.  
 PA (GERT) GENENTECH-INC.  
 PI Bass SH, Cunningham BC, Fuh G, Lowman HB, Matthews DJ;  
 DR WPI; 92-096838/12.  
 PT New method of modifying polypeptide hormone-receptor complex - to  
 PT produce human growth hormone variant, useful for stimulating  
 PT lactogenic and somatogenic response  
 PS Disclosure; Page 41; 74pp; English.  
 CC This truncated human growth hormone binding protein was used as the  
 CC basis of mutagenesis experiments to generate receptor variants. The  
 CC results obtained with the shorter receptor were indistinguishable  
 CC from those obtained with the 246 amino acid hGHbp. The sequence was  
 CC modified at the metal binding site. Modification of metal ion  
 CC chelation allows the response of a cell, organ or animal to a  
 CC hormone to be controlled, e.g. Zn ion binding is required for tight  
 CC binding of hGH to the prolactin receptor but not to the hGH  
 CC receptor. See e.g. R24274 for an example of a preferred variant.  
 SQ Sequence 237 AA;

Thu May 11 06:49:34 2000

US-09-376-430-2-05.rag

Page 7

Query Match 7.2%; Score 111; DB 1; Length 237;  
Best Local Similarity 33.3%; Pred. No. 3.60e-01;  
Matches 17; Conservative 11; Mismatches 22; Indels 1; Gaps 1;  
Db 172 LEYELQYKEVNETKWKMDPILTTSPVYSLKVDKEYEVVRVRSKQRNSGYG 222  
QY 146 LLYEVQYRSPFDEWQSKQEN-TCNVITIEGLDAEKCYSEFWVRVKAMEDVYG 195

Search completed: Wed May 10 11:44:50 2000  
Job time : 13 secs.





Thu May 11 06:49:34 2000

\*\*\*\*\*  
Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd  
\*\*\*\*\*  
MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:45:07 2000; MasPar time 48.31 Seconds  
Tabular output not generated.  
\*\*\*\*\*  
(TM)  
\*\*\*\*\*

Title: >US-09-376-430-2  
Description: (23-225) from US09376430A.pap (5 of 25)  
Perfect Score: 1545  
Sequence: 1 OGGNAEGVQIIYFNLETV.....VTCWQGEIRDACAETPTTP 203  
Scoring table: PAM 150  
Gap 11  
Searched: 131253 seqs, 12956647 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: a-issued  
1:5A\_COMB 2:5B\_COMB 3:PCT\_COMB 4:backfiles1  
Statistics: Mean 29.992; Variance 124.519; scale 0.241  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.  
SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	195	12.6	230	1	US-08-595- Sequence 11, Applicati	1.39e-08
2	195	12.6	230	1	US-08-052- Sequence 11, Applicati	1.39e-08
3	195	12.6	252	1	US-08-595- Sequence 9, Applicati	1.39e-08
4	195	12.6	252	1	US-08-052- Sequence 9, Applicati	1.39e-08
5	195	12.6	347	1	US-08-595- Sequence 7, Applicati	1.39e-08
6	195	12.6	347	1	US-08-052- Sequence 7, Applicati	1.39e-08
7	195	12.6	369	1	US-08-595- Sequence 4, Applicati	1.39e-08
8	195	12.6	369	1	US-08-052- Sequence 4, Applicati	1.39e-08
9	188	12.2	369	3	PCT-US94-0 Sequence 69, Applicati	6.00e-08
10	188	12.2	369	2	US-08-424- Sequence 2, Applicati	6.00e-08
11	130	8.4	383	1	US-08-609- Sequence 2, Applicati	1.96e-01
12	113	7.3	897	1	US-07-960- Sequence 2, Applicati	1.96e-01
13	99	6.4	982	2	US-08-609- Sequence 4, Applicati	7.64e+00
14	93	6.0	380	1	US-08-609- Sequence 4, Applicati	7.64e+00
15	89	5.8	446	1	US-08-485- Sequence 2, Applicati	1.58e+01
16	89	5.8	446	2	US-09-003- Sequence 2, Applicati	1.58e+01
17	88	5.7	933	1	US-08-162- Sequence 16, Applicati	2.19e+01
18	87	5.6	488	2	US-08-599- Sequence 5, Applicati	2.19e+01
19	87	5.6	658	2	US-08-825- Sequence 2, Applicati	2.19e+01
20	87	5.6	708	1	US-08-308- Sequence 2, Applicati	2.19e+01
21	87	5.6	708	2	US-08-059- Sequence 2, Applicati	2.19e+01
22	87	5.6	708	1	US-07-797- Sequence 2, Applicati	2.19e+01
23	87	5.6	708	2	US-09-058- Sequence 2, Applicati	2.19e+01

24	87	5.6	708	3	PCT-US95-0 Sequence 2, Applicatio	2.19e+01
25	87	5.6	918	2	US-08-825- Sequence 6, Applicatio	2.19e+01
26	85	5.5	288	1	US-08-375- Sequence 9, Applicatio	3.10e+01
27	85	5.5	288	1	US-08-752- Sequence 9, Applicatio	3.10e+01
28	83	5.4	187	1	US-08-049- Sequence 33, Applicati	4.36e+01
29	83	5.4	187	1	US-08-312- Sequence 33, Applicati	4.36e+01
30	83	5.4	187	3	PCT-US94-0 Sequence 33, Applicati	4.36e+01
31	83	5.4	574	1	US-08-049- Sequence 2, Applicatio	4.36e+01
32	83	5.4	574	1	US-08-312- Sequence 2, Applicatio	4.36e+01
33	83	5.4	574	3	PCT-US94-0 Sequence 2, Applicatio	4.36e+01
34	83	5.4	1447	3	PCT-US95-0 Sequence 141, Applicat	5.17e+01
35	82	5.3	128	3	PCT-US95-0 Sequence 2, Applicati	5.17e+01
36	82	5.3	541	2	US-08-702- Sequence 11, Applicati	5.17e+01
37	82	5.3	970	2	US-08-449- Sequence 11, Applicati	5.17e+01
38	82	5.3	970	3	PCT-US95-0 Sequence 11, Applicati	5.17e+01
39	82	5.3	970	3	PCT-US95-0 Sequence 15, Applicati	7.24e+01
40	80	5.2	607	2	US-08-878- Sequence 101, Applicat	6.12e+01
41	81	5.2	948	2	US-08-469- Sequence 5, Applicatio	6.12e+01
42	81	5.2	995	2	US-08-673- Sequence 18, Applicati	6.12e+01
43	81	5.2	995	1	US-08-162- Sequence 12, Applicati	6.12e+01
44	81	5.2	1011	1	US-08-162- Sequence 5, Applicatio	8.57e+01
45	79	5.1	1911	2	US-08-800- Sequence 5, Applicatio	8.57e+01

ALIGNMENTS

RESULT 1  
ID US-08-595-974-11 STANDARD; PRT; 230 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
XX  
DE  
XX  
XX  
Sequence 11, Application US/08595974  
Sequence 11, Application US/08595974  
Patent No. 5705608  
GENERAL INFORMATION:  
APPLICANT: SUGAMURA, KAZUO  
APPLICANT: TAKESHITA, TOSHIKAZU  
APPLICANT: ASANO, HIRONOBU  
APPLICANT: NAKAMURA, MASATAKA  
APPLICANT: SHIMAMURA, TOSHIRO  
APPLICANT: SUZUKI, MANABU  
APPLICANT: HAMURO, JUNJI  
TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/595,974  
APPLICATION NUMBER: US/08/595,974  
FILING DATE: 06-FEB-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/052,205  
FILING DATE: 22-APR-1993  
APPLICATION NUMBER: JP 104947/1992  
FILING DATE: 23-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5705608man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 10-615-0X

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/052,205  
FILING DATE: 19930422  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 104947/1992

```
CC: SOFTWARE: PatentIn Release #1  
CC: OPERATING SYSTEM: PC-DOS/MS-DOS  
CC: CURRENT APPLICATION DATA:
```



CC ADDRESSEE: P.C.  
 CC STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 CC CITY: Arlington  
 CC STATE: Virginia  
 CC COUNTRY: U.S.A.  
 CC ZIP: 22202  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patentin Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: JP 104947/1992  
 CC FILING DATE: 23-APR-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Oblon, No. 5705608man F.  
 CC REGISTRATION NUMBER: 24,618  
 CC REFERENCE/DOCKET NUMBER: 10-615-0X  
 CC TELEPHONE: (703) 413-3000  
 CC TELEFAX: (703) 413-3000  
 CC INFORMATION FOR SEQ ID NO: 7:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 347 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 347 AA; 39920 MW; 666553 CN;

Query Match 12.6%; Score 195; DB 1; Length 347;  
 Best Local Similarity 25.9%; Pred. No. 1.39e-08;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 37 EVQCFVFNVEYMNCTNSSEPOPTNLTHWYKNSDNDKVKQSHYLFSEETSGCQLQ 96  
 QY 31 QIQIIFNLETVQVWNASKYSR-TNLTFFHYRF-NGD-EAYDQCTNLYLQEGHTSGCLLD 87  
 Db 97 KKEIHLVQTFVVOQLQDPREPRQATQMLKQLNVLIPWAPENLTLHKLSESOLELNNRNF 156  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRWV-VYIL-KPSSPKHVRF-SWHQDAVTVTCSDL 142  
 Db 157 LNHCLHLVQYRTDWDHSHWTEQSDYVRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 215  
 QY 143 YGD-LLYEVQYRSFPDTEW-OSKQENTCNVTIEGLDAEKCYSEFWVRVKAMEDVYGPDTYP 200  
 Db 216 SEWSHPH 224  
 QY 201 SDWSEVTCW 209

RESULT 6  
 ID US-08-052-205-7 STANDARD; PRT; 347 AA.  
 XX  
 AC xxxxxx  
 DT  
 XX  
 XX  
 XX  
 XX  
 DE Sequence 7, Application US/08052205  
 CC Sequence 7, Application US/08052205  
 CC Patent No. 5510259  
 CC GENERAL INFORMATION:  
 CC APPLICANT: SUGAMURA, KAZUO  
 CC APPLICANT: TAKESHITA, TOSHIKAZU  
 CC APPLICANT: ASAO, HIRONOBU  
 CC APPLICANT: NAKAMURA, MASATAKA

CC APPLICANT: SHIMAMURA, TOSHIRO  
 CC APPLICANT: SUZUKI, MANABU  
 CC TITLE OF INVENTION: HUMAN IL-2 RECEPTOR GAMMA CHAIN MOLECULE  
 CC NUMBER OF SEQUENCES: 21  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: P.C.  
 CC STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 CC CITY: Arlington  
 CC STATE: Virginia  
 CC COUNTRY: U.S.A.  
 CC ZIP: 22202  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patentin Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/052,205  
 CC FILING DATE: 19930422  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: JP 104947/1992  
 CC FILING DATE: 23-APR-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Oblon, No. 5510259man F.  
 CC REGISTRATION NUMBER: 24,618  
 CC REFERENCE/DOCKET NUMBER: 10-615-0X  
 CC TELEPHONE: (703) 413-3000  
 CC TELEFAX: (703) 413-2220  
 CC INFORMATION FOR SEQ ID NO: 7:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 347 amino acids  
 CC TYPE: AMINO ACID  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 347 AA; 39920 MW; 666553 CN;

Query Match 12.6%; Score 195; DB 1; Length 347;  
 Best Local Similarity 25.9%; Pred. No. 1.39e-08;  
 Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
 Db 37 EVQCFVFNVEYMNCTNSSEPOPTNLTHWYKNSDNDKVKQSHYLFSEETSGCQLQ 96  
 QY 31 QIQIIFNLETVQVWNASKYSR-TNLTFFHYRF-NGD-EAYDQCTNLYLQEGHTSGCLLD 87  
 Db 97 KKEIHLVQTFVVOQLQDPREPRQATQMLKQLNVLIPWAPENLTLHKLSESOLELNNRNF 156  
 QY 88 AEQRDDILYFSIR-NGTH-PVFTASRWV-VYIL-KPSSPKHVRF-SWHQDAVTVTCSDL 142  
 Db 157 LNHCLHLVQYRTDWDHSHWTEQSDYVRHKFSLPSVDGQKRYTFVRSR-FNPLCGSAQHW 215  
 QY 143 YGD-LLYEVQYRSFPDTEW-OSKQENTCNVTIEGLDAEKCYSEFWVRVKAMEDVYGPDTYP 200  
 Db 216 SEWSHPH 224  
 QY 201 SDWSEVTCW 209

RESULT 7  
 ID US-08-595-974-4 STANDARD; PRT; 369 AA.  
 XX  
 AC xxxxxx  
 DT  
 XX  
 XX  
 DE Sequence 4, Application US/08595974  
 CC Sequence 4, Application US/08595974  
 CC Patent No. 5705608





US-09-376-430-2-05.rai

Thu May 11 06:49:34 2000

Db 119 KEDTQYQTEVQVQDQKQKQRAVQKLNQNLQVQIPRAPENLTLNSLSESOLELRWKSRR 178  
QY 88 AEQRDDLYFSIR-NGTH-PVFTASRW-VVYL-KPSSPKHVRFS-WHQDAVTV--TCSD 140  
Db 179 IKERCQYLYQYRSNRDRSTELVNEHPRFSLPSVDELKRYTFRVRSRY-NPICGSSQ 237  
QY 141 LSYGDLLEYQYRSPEDTEW-QSKOENTCNVTIEGLDAEKCYSEFWVRVKAMEDVYGPDTY 199  
Db 238 WSKWSQPVHW 247  
QY 200 PSDSEVTCW 209

RESULT 11 US-08-609-572-2 STANDARD; PRT; 383 AA.

Sequence 2, Application US/08609572  
Patent No. 5710023  
GENERAL INFORMATION:  
APPLICANT: Collins, Mary  
APPLICANT: Donaldson, Debra  
APPLICANT: Fitz, Lori  
APPLICANT: Neben, Tamlyn  
APPLICANT: Whitters, Matthew  
APPLICANT: Wood, Clive  
TITLE OF INVENTION: CYTOKINE RECEPTOR CHAIN  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: US/08/609,572  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: G15268  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 383 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
FEATURE: INFORMATION: Human GM-CSF receptor; Signal Sequence: -17 to -1;  
SEQUENCE 383 AA; 44882 MW; 795803 CN;

Query Match 8.48; Score 130; DB 1; Length 383;  
Best Local Similarity 30.28; Pred. No. 7.55e-03;  
Matches 26; Conservative 22; Mismatches 32; Indels 6; Gaps 5;

Db 127 EGSLETKIQDMKCIYYNQYLVCSWPKCKYVSDNTYMEFWYEGLDHAL-QCADIYLOHD 185  
QY 23 QGGAAGVQ-IOIIFYNLETFVQVTWNSK--YSRNLTFHFRNG-DEADYDQCTNYLLOE 78

Db 186 EKNVGCKLSNLDSSDYKDFEICYNGS 211  
QY 79 GHTSGCLLDQAEQRDDIL-YFSIRNGT 103  
RESULT 12 US-07-960-389-2 STANDARD; PRT; 897 AA.  
Sequence 2, Application US/07960389  
Patent No. 5705611  
GENERAL INFORMATION:  
APPLICANT: HAYASHIDA, Kasuhiro;  
TITLE OF INVENTION: Human GM-CSF Receptor Component  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schering-Plough Corporation  
STREET: 2000 Galloping Hill Road  
CITY: Kenilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disc  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: System Software 7.1  
SOFTWARE: Macintosh Word 5.1a  
CURRENT APPLICATION DATA:  
FILING DATE: 07-JAN-1993  
CLASSIFICATION: 800  
PRIOR APPLICATION NUMBER: 554,745  
FILING DATE: 18-JUL-1990  
APPLICATION NUMBER: PCT/US 91/04846  
FILING DATE: 16-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Blasdale, John H. C.  
REGISTRATION NUMBER: 31,895  
REFERENCE/DOCKET NUMBER: DX01430  
TELEPHONE: (908) 298-2902  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 897 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE: INFORMATION: Human GM-CSF receptor; Signal Sequence: -17 to -1;  
SEQUENCE 897 AA; 97323 MW; 4443497 CN;

Query Match 7.38; Score 113; DB 1; Length 897;  
Best Local Similarity 31.9%; Pred. No. 1.96e-01;  
Matches 22; Conservative 10; Mismatches 34; Indels 3; Gaps 3;

Db 372 FEIQYKDKAT-WKDSKTELTQNAHSMALPALEPSTRYVARVRVRSYTCYNGIWSNSE 430  
QY 148 YEQYASQEDTEW-QSKOENTCNVTIEGLDA-EKCYSEFWVRVKAMEDVYGPDTYPSDSE 205  
Db 431 ARSWDFESV 439  
QY 206 VTCWQGEI 214

RESULT 13





Thu May 11 06:49:34 2000

CC NUMBER OF SEQUENCES: 2  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Merchant & Gould  
 CC STREET: 3100 No. 569560west Center, 90 S. 7th Street  
 CC CITY: Minneapolis  
 CC STATE: Minnesota  
 CC COUNTRY: U.S.A.  
 CC ZIP: 55402  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/485,216  
 CC FILING DATE: 07-JUN-1995  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: WO 94/26907  
 CC FILING DATE: 13-MAY-1994  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Mueller, Douglas P.  
 CC REGISTRATION NUMBER: 30,300  
 CC REFERENCE/DOCKET NUMBER: 7933.29US02  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 612/332-5300  
 CC TELEFAX: 612/332-9081  
 CC INFORMATION FOR SEQ ID NO: 2:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 446 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 446 AA; 50277 MW; 986088 CN;

Query Match 5.8%; Score 89; DB 1; Length 446;  
 Best Local Similarity 31.3%; Pred. No. 1.55e+01;  
 Matches 10; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

Db 6 RIPLLYVHLTMGSKWSVDKTHRTFLGFIYIF 37  
 QY 31 QIQIIFNLETQVQVWNASKYSRTNLTFFHYRF 62

Search completed: Wed May 10 11:46:03 2000  
 Job time : 56 secs.



\*\*\*\*\*  
 M P S R L H  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 11:44:00 2000; MasPar time 15.07 Seconds  
 Tabular output not generated.  
 635.443 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (23-225) from US09376430A.pep (5 of 25)  
 Perfect Score: 1545  
 Sequence: 1 QGGAEGVQIQIIFNLETV.....VTCWQGEIRDCAETPTTP 203

Scoring table: PAM 150  
 Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: pir62  
 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 42.929; Variance 76.795; scale 0.559

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	210	13.6	373	2	A55718 Interleukin-2 recepto	2.83e-21
2	195	12.6	369	2	A42565 Interleukin-2 recepto	2.29e-18
3	188	12.2	369	2	I49280 Interleukin-2 recepto	5.00e-17
4	115	7.4	878	1	A40091 Interleukin-3 recepto	3.92e-04
5	113	7.3	897	1	A39255 cytokine receptor com	8.09e-04
6	109	7.1	435	2	E71350 probable aspartate am	3.37e-03
7	108	7.0	896	1	A35782 cytokine receptor com	4.79e-03
8	100	6.5	638	2	S04530 somatotropin receptor	7.48e-02
9	100	6.5	638	2	A33991 somatotropin receptor	7.48e-02
10	99	6.4	4436	2	E71086 hypothetical protein	1.05e-01
11	99	6.4	6839	2	S57242 twitchin - Caenorhabd	1.05e-01
12	97	6.3	302	2	S50579 hypothethical protein	2.03e-01
13	97	6.3	311	2	S77803 hypothethical protein	2.03e-01
14	96	6.2	638	2	S12136 somatotropin receptor	2.81e-01
15	96	6.2	896	2	I56563 Interleukin-3 recepto	2.81e-01
16	94	6.1	638	2	B28176 somatotropin receptor	5.38e-01
17	94	6.1	26926	1	I38344 titin, cardiac muscle	5.38e-01
18	93	6.0	173	2	T15695 hypothethical protein	7.41e-01
19	92	6.0	358	2	T07719 aldose 1-epimerase-li	1.02e-00
20	92	6.0	361	2	JN0716 glutamate--ammonia li	1.02e-00
21	92	6.0	684	2	S60266 novel antigen recepto	1.02e-00
22	89	5.8	255	2	E72226 conserved hypothetica	2.61e+00
23	90	5.8	376	2	D72493 hypothethical protein	1.91e+00

24	90	5.8	629	2	A30001 ribosomal protein S6	1.91e+00
25	89	5.8	1286	2	A42150 p-glycoprotein pgp1 -	2.61e+00
26	88	5.7	137	2	JQ2325 virion protein - Afri	3.55e+00
27	88	5.7	562	1	O0BEH5 phosphotransferase (E	3.55e+00
28	88	5.7	935	1	I41193 outer membrane protei	3.55e+00
29	88	5.7	983	2	B45583 receptor tyrosine kin	3.55e+00
30	86	5.6	43	2	T06456 myb-like protein - ga	6.51e+00
31	87	5.6	418	1	BVEBHA glutamyl--trNA reducta	4.81e+00
32	87	5.6	430	2	JC2334 lanosterol 14alpha-de	4.81e+00
33	87	5.6	486	2	JC4240 lanosterol 14alpha-de	4.81e+00
34	87	5.6	503	2	JC4758 lanosterol 14alpha-de	4.81e+00
35	86	5.6	632	1	JC4919 2-oxoacid--ferredoxin	6.51e+00
36	87	5.6	634	2	S33339 somatotropin receptor	4.81e+00
37	86	5.6	732	2	A72118 exodeoxyribonuclease	6.51e+00
38	87	5.6	918	2	A36337 membrane glycoprotein	4.81e+00
39	87	5.6	939	2	I41197 eae protein (enteropa	4.81e+00
40	87	5.6	1354	2	I13930 tripeptidyl-peptidase	4.81e+00
41	85	5.5	129	2	S03534 Ig heavy chain precur	8.79e+00
42	85	5.5	471	2	JC1403 glutamate--ammonia li	8.79e+00
43	85	5.5	706	2	S42174 NTPase - African swin	8.79e+00
44	85	5.5	706	2	JQ2210 probable nucleic acid	8.79e+00
45	85	5.5	6658	2	T13931 projectin - fruit fly	8.79e+00

## ALIGNMENTS

RESULT 1  
 ENTRY A55718 #type complete  
 TITLE interleukin-2 receptor gamma chain precursor - dog  
 ORGNAME #formal\_name Canis lupus familiaris #common\_name dog  
 DATE 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 23-Jul-1999  
 A55718  
 A55718  
 Henthorn, P.S.; Somberg, R.L.; Fimiani, V.M.; Puck, J.M.;  
 Patterson, D.F.; Felsburg, P.J.  
 Genomics (1994) 23:69-74  
 IL-2Rgamma gene microdeletion demonstrates that canine  
 X-linked severe combined immunodeficiency is a homologue of  
 the human disease.  
 #cross-references MIM:95130114  
 #accession A55718  
 #status preliminary  
 #molecule\_type mRNA  
 #residues 1-373 #label HEN  
 #cross-references GB:U04361; NID:G517411; PID:AA048403.1; PID:G517412  
 CLASSIFICATION #superfamily interleukin-2 receptor gamma chain  
 KEYWORDS cytokine receptor; duplication  
 SUMMARY #length 373 #molecular-weight 42516 #checksum 7960

Query Match	13.6%	Score	210	DB	2	Length	373
Best Local Similarity	25.9%	Pred. No.	2.83e-21				
Matches	49	Conservative	56	Mismatches	73	Indels	11
Gaps	10						
Db	59	EVQCFVFNVMCTWNSSEPRPTNLTLYWYKNSNDKQVEGCHYLFSEVNTAGCWLQ	118				
QY	31	QIQIIFNLETVQVWNASKYSR-TNLTIFYRF-NGDEA-YDQCTNYLLQEGHSGCLLD	87				
Db	119	KEIHLIYETVWLRDPRPRQSTQKLNQVWPAPENLTILNLSQELSLWSNRH	178				
QY	88	AEQRD--DILYFSIRNGTHPVFTAS-RWMVYVL-KPSSPKHVRP-SWHQDAVTTCSDLS	142				
Db	179	LDRCLEHVVOYRSDWRSTEQSVDRHSFSLPSVDQKQFYTFYFVRSRY-NPLCGSAQRW	237				
QY	143	YGD-LLVEVOYRSPFDTEW-QSKQENTCNVTIEGLDAEKCYSEWVRKAMEDVYGPDTYP	200				
Db	238	SEWSHPHWH 246					
QY	201	SDWSEVTCW 209					
RESULT 2							
ENTRY							



```

##status      preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues    1-369 #label RES
##cross-references GB:575852; NID:g861554; PIDN:AAB32904.1; PID:g861555
GENETICS
#gene         IL-28gamma
#exons        39/1, 90/2; 152/1; 199/3; 254/1; 286/2; 308/3
#introns      The high affinity receptor is a heterotrimer of alpha (see
#complex      PIR:UHS2), beta (see PIR:A35052), and gamma chains;
#              heterodimers of alpha or beta and gamma chains are
#              intermediate affinity receptors.
FUNCTION
#description   receptor for interleukin-2
#pathway       Interleukin-2 stimulated growth and differentiation of T
#              cells, B cells, NK cells, LAK cells, monocytes,
#              macrophages, and oligodendrocytes
CLASSIFICATION
#superfamily  #superfamily interleukin-2 receptor gamma chain
#cytokine     cytokine receptor; duplication; glycoprotein; transmembrane
#protein      protein
FEATURE
1-22          #domain signal sequence #status predicted #label SIG\
23-369        #product interleukin-2 receptor gamma chain #status
#predicted #label MAT\
256-284       #domain transmembrane #status predicted #label TMM\
71,75,84,96,159, #binding site carbohydrate (Asn) (covalent) #status
164,306       Predicted
SUMMARY       #length 369 #molecular-weight 42241 #checksum 6734

Query Match      12.28; Score 188; DB 2; Length 369;
Best Local Similarity 26.88; Pred. No. 5.00e-17;
Matches 51; Conservative 52; Mismatches 75; Indels 12; Gaps 11;

Db 59 EVOCFVFNEMNCTNSSEPOATNLTHRYKYVSDNNTFQCSHYLPKETSQCIQ 118
QY 31 QIQIIFNLETVOVTWNAKYSR-TNLTFHYRNF-GDE-AYDOCTNYLQEGHTSGCLLD 87
Db 119 KEDIQLYQTFVVOQLPQKPRRAVKQLNQLNIPAPENLTLSNLSQLELRWKSRI 178
QY 88 AEQRDDILYFSIR-NGTH-PVFTASRWV-VYVL-KPSSPKHVRFS-WHODAVTV--TCSD 140
Db 179 IKERCQLYQVYRSNRDRSTELIVNHEPRFSLPSVDLKYFRFRVRSRY-NPICSSQ 237
QY 141 LSGYDLLEYQVYRSPDTEW-QSKQENTCNVTIEGLDAERKCYSFVVRVKAMEDVYGPDTY 199
Db 238 WSKWSQPVHW 247
QY 200 PSDWSEVTCW 209

RESULT 4
ENTRY      A40091 #type complete
TITLE      Interleukin-3 receptor beta chain precursor - mouse
ORGANISM   Mus musculus #common_name mouse
DATE       30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change
22-Jun-1999
ACCESSIONS A40091; A43022
REFERENCE   Itoh, N.; Yonehara, S.; Schreurs, J.; Gorman, D.M.; Maruyama
#authors    K.; Ishii, A.; Yahara, I.; Arai, K.; Miyajima, A.
#journal    Science (1990) 247:324-327
#title      Cloning of an interleukin-3 receptor gene: a member of a
#            distinct receptor gene family.
#cross-references MUID:90117145
#accession  A40091
#status     nucleic acid sequence not shown
##molecule_type mRNA
##residues  1-878 #label ITO
REFERENCE   A43022
#authors    Gorman, D.M.
#submitted  submitted to GenBank, November 1989
#accession  A43022
#molecule_type mRNA

```

```

##residues    1-815,'Q',817-878 #label GOR
##cross-references GB:M29855; NID:g198342; PIDN:AAA3295.1; PID:g309406
COMMENT      In mice there are two classes of high-affinity IL-3 receptors. One
              contains this IL-3-specific beta chain and the other contains the
              beta chain also shared by high-affinity IL-5 and GM-CSF
              receptors.
CLASSIFICATION
#superfamily  #superfamily interleukin-3 receptor beta chain; cytokine
#receptor     receptor homology
#cytokine     cytokine receptor; duplication; transmembrane protein
#product      #product interleukin-3 receptor beta chain #status
#predicted #label MAT\
23-878        #domain signal sequence #status predicted #label EXT\
23-440        #domain extracellular #status predicted #label CRS1\
39-236        #domain cytokine receptor homology #label CRS1\
234-433       #domain cytokine receptor homology #label CRS2\
441-452       #domain transmembrane #status predicted #label TMM\
463-878       #domain intracellular #status predicted #label INT\
SUMMARY       #length 878 #molecular-weight 97222 #checksum 2346

Query Match      7.48; Score 115; DB 1; Length 878;
Best Local Similarity 32.48; Pred. No. 3.92e-04;
Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 370 YIDHTFQVQYKKSES-WKDSKTENLGRVNSMDLPDPTSYCARVRKPIDSD-YD-GI 426
QY 143 YGDLLEYQVYRSPDTEW-QSKQENTCNV-TIE-G-LDAEKCYSFVVRVKAMEDVYGPDT 198
Db 427 W-SEWSNEYTW 436
QY 199 YPSDWSEVTCW 209

RESULT 5
ENTRY      A39255 #type complete
TITLE      cytokine receptor common beta chain precursor - human
ORGANISM   Homo sapiens #common_name man
DATE       30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change
02-Sep-1997
ACCESSIONS A39255
REFERENCE   Hayashida, K.; Kitamura, T.; Gorman, D.M.; Arai, K.; Yokota,
#authors    T.; Miyajima, A.
#journal    Proc. Natl. Acad. Sci. U.S.A. (1990) 87:9655-9659
#title      Molecular cloning of a second subunit of the receptor for
#            human granulocyte-macrophage colony-stimulating factor
#            (GM-CSF): reconstruction of a high-affinity GM-CSF
#            receptor.
#cross-references MUID:91088571
#accession  A39255
#molecule_type mRNA
##residues  1-897 #label HAY
##cross-references GB:M38275
COMMENT      The human high-affinity IL-3, IL-5, and GM-CSF receptors have
              ligand-specific alpha chains and share this common beta chain.
GENETICS
#gene       GDB:CSF2RB
#cross-references GDB:126838; OMIM:138981
#map_position 22q13.1-22q13.1
CLASSIFICATION
#superfamily  #superfamily interleukin-3 receptor beta chain; cytokine
#receptor     receptor homology
#cytokine     cytokine receptor; duplication;
#protein      transmembrane protein
FEATURE
1-16         #domain signal sequence #status predicted #label SIG\
17-897       #product cytokine receptor common beta chain #status
#predicted #label MAT\
17-443       #domain extracellular #status predicted #label EXT\
35-232       #domain cytokine receptor homology #label CRS1\
250-431      #domain cytokine receptor homology #label CRS2\
444-460      #domain transmembrane #status predicted #label TMM\
461-897      #domain intracellular #status predicted #label INT\
SUMMARY       #length 897 #molecular-weight 97323 #checksum 753

```

```

Query Match      7.3%; Score 113; DB 1; Length 897;
Best Local Similarity 31.9%; Pred. No. 8.09e-04;
Matches 22; Conservative 10; Mismatches 34; Indels 3; Gaps 3;

Db 372 FEIQYKDTAT-WKDSKTEILONAHSMALPALESTRYWARVRVTSYTGNGYIWSWSE 430
QY 148 YEVOYRSPFTTEW-OSKOENTCNVTIEGLDA-EKCYSFVVRVKAMEDYVDPYPSDSE 205
Db 431 ARSWDTESV 439
QY 206 VTCWQGEI 214

RESULT 6
ENTRY #type complete
TITLE probable aspartate aminotransferase (tpaat) - syphilis
ORGANISM #formal_name Treponema pallidum subsp. pallidum #common_name
        syphilis spirochete
DATE 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change
ACCESSIONS E71350
REFERENCE E71350
#authors Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.;
        Sutton, G.G.; Dodson, R.; Gwinn, M.; Hickey, E.K.; Clayton,
        R.; Ketchum, K.A.; Sodergren, E.; Hardham, J.M.; McLeod,
        M.P.; Salzberg, S.; Peterson, J.; Khalak, H.; Richardson,
        D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDonald,
        L.; Artiach, P.; Bowman, C.; Cotton, M.D.; Fujii, C.;
        Garland, S.; Hatch, B.; Horst, K.; Roberts, K.; Watthey,
        L.; Weidman, J.; Smith, H.O.; Venter, J.C.
#journal Science (1998) 281:375-388
#title Complete genome sequence of Treponema pallidum, the syphilis
        spirochete.
#cross-references MUID:98332770
#accession E71350
#status preliminary; nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
##residues 1-435 #label COL
##cross-references GB:AE001204; GB:AE000520; NID:g3322492; PID:g3322493
#experimental_source strain Nichols
GENETICS
#gene TP0223
#summary #length 435 #molecular-weight 47304 #checksum 4156

Query Match      7.1%; Score 109; DB 2; Length 435;
Best Local Similarity 35.1%; Pred. No. 3.37e-03;
Matches 20; Conservative 11; Mismatches 21; Indels 5; Gaps 5;

Db 206 AVLVICDD-AYSGFEYASIMRGSFFARFAQAHNKCAIKDGLTKEE-YA-WGLRV 259
QY 133 AVTVTCSDLSYGLLYEVOY-RSPFDTEWQSKOENTCNVTIEGLDAEKCYSFV-VRV 187

RESULT 7
ENTRY #type complete
TITLE cytokine receptor common beta chain precursor - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change
ACCESSIONS A35782
REFERENCE A35782
#authors Gorman, D.M.; Itoh, N.; Kitamura, T.; Schreurs, J.; Yonehara,
        S.; Yahara, I.; Arai, K.; Miyajima, A.
#journal Proc. Natl. Acad. Sci. U.S.A. (1990) 87:5459-5463
#title Cloning and expression of a gene encoding an interleukin 3
        receptor-like protein: identification of another member of
        the cytokine receptor gene family.
#cross-references MUID:90319131
#accession A35782
#molecule_type mRNA

```

```

##residues 1-896 #label GOR
##cross-references GB:M34397; NID:g191821; PIDN:AAA37204.1; PID:g309101
COMMENT Mouse high-affinity IL-5, GM-CSF, and one class of high-affinity
        IL-3 receptors have ligand-specific alpha chains and share this
        common beta chain.
CLASSIFICATION #superfamily interleukin-3 receptor beta chain; cytokine
        receptor homology
KEYWORDS cytokine receptor; duplication; transmembrane protein
FEATURE
1-22 #domain signal sequence #status predicted #label SIG\
23-896 #product cytokine receptor common beta chain #status
        predicted #label MAT\
23-441 #domain extracellular #status predicted #label EXT\
39-235 #domain cytokine receptor homology #label CRS1\
253-434 #domain cytokine receptor homology #label CRS2\
442-463 #domain transmembrane #status predicted #label TMM\
464-896 #domain intracellular #status predicted #label INT\
SUMMARY #length 896 #molecular-weight 99110 #checksum 6643

Query Match      7.0%; Score 108; DB 1; Length 896;
Best Local Similarity 30.6%; Pred. No. 4.79e-03;
Matches 22; Conservative 18; Mismatches 24; Indels 8; Gaps 8;

Db 370 SFIEHTFOVQYKKKSDS-WEDSKTENLDRAHSMDSQLEPDTSCARVRVKPISN-YD-G 426
QY 142 SYGDLLEYQVRSPPDTEWQ-SKQEN-TCNVITIE-G-LDAEKCYSFVVRVKAMEDYVGP 197
Db 427 IW-SKWSEYTW 437
QY 198 TYPSONSEVTCW 209

RESULT 8
ENTRY #type complete
TITLE somatotropin receptor, hepatic precursor - human
ALTERNATE_NAMES growth hormone receptor
CONTAINS somatotropin-binding protein, serum
ORGANISM #formal_name Homo sapiens #common_name man
DATE 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
ACCESSIONS S04530
REFERENCE S04530
#authors Leung, D.W.; Spencer, S.A.; Cachianes, G.; Hammonds, R.G.;
        Collins, C.; Henzel, W.J.; Barnard, R.; Waters, M.J.; Wood,
        W.I.
#journal Nature (1987) 330:537-543
#title Growth hormone receptor and serum binding protein:
        purification, cloning and expression.
#cross-references MUID:88065896
#accession S04530
#molecule_type mRNA
##residues 1-638 #label LEU
##cross-references EMBL:X06562; NID:g31737; PID:g31738
CLASSIFICATION #superfamily fibronectin type III repeat homology
KEYWORDS liver; transmembrane protein
FEATURE
1-18 #domain signal sequence #status predicted #label SIG\
19-638 #product somatotropin receptor #status predicted #label
        MAT\
265-288 #domain transmembrane #status predicted #label TMM
SUMMARY #length 638 #molecular-weight 71499 #checksum 4388

Query Match      6.5%; Score 100; DB 2; Length 638;
Best Local Similarity 32.7%; Pred. No. 7.48e-02;
Matches 17; Conservative 11; Mismatches 22; Indels 2; Gaps 2;

Db 190 LEVELQYKEVNEFKWMDILITSPVYSLKVDKEVVRVRSQRQNSGNYG 241
QY 146 LLVEVQVRSPPDTEWQSKQEN-TCNVITIEGLDAEKCYSFVVRVKAMEDV-YG 195

RESULT 9
ENTRY #type complete

```

```

TITLE
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
#journal
#title
#cross-references
#accession
#status
#molecule_type
#residues
#cross-references
GENETICS
#gene
#map_position
#classification
#keywords
#summary
Query Match
Best Local Similarity
Matches
Db
Qy
RESULT
ENTRY
TITLE
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
#journal
#title
#cross-references
#accession
#status
#molecule_type
#residues
#cross-references
#note
GENETICS
#gene
#summary
Query Match
Best Local Similarity

```

somatotropin receptor precursor - human  
 #formal\_name Homo sapiens #common\_name man  
 23-Mar-1990 #sequence\_revision 23-Mar-1990 #text\_change  
 24-Sep-1998  
 A33991  
 Godowski, P.J.; Leung, D.W.; Meacham, L.R.; Galgani, J.P.;  
 Hellmiss, R.; Keret, R.; Rotwein, P.S.; Parks, J.S.; Laron,  
 Z.; Wood, W.I.  
 Proc. Natl. Acad. Sci. U.S.A. (1989) 86:8083-8087  
 Characterization of the human growth hormone receptor gene  
 and demonstration of a partial gene deletion in two  
 patients with Laron-type dwarfism.  
 #cross-references MUID:90046742  
 A33991 preliminary  
 #status  
 #molecule\_type DNA  
 #residues 1-638 #label GOD  
 #cross-references GB:M28458; GB:M28459; GB:M28460; GB:M28461;  
 GB:M28462; GB:M28463; GB:M28464; GB:M28465;  
 GB:M28466; NID:g183168; PID:g183171  
 GENETICS  
 #gene GDB:GHR  
 #map\_position Sp13-Sp12  
 #classification #superfamily fibronectin type III repeat homology  
 #keywords transmembrane protein  
 #summary #length 638 #molecular\_weight 71499 #checksum 4481  
 Query Match 6.5%; Score 100; DB 2; Length 638;  
 Best Local Similarity 32.7%; Pred. No. 7.48e-02;  
 Matches 17; Conservative 11; Mismatches 22; Indels 2; Gaps 2;  
 Db 190 LEVELYQKVENETKWKMDPILITSPVYSLKVDKEYEYVRKSKQNSNGY 241  
 Qy 146 LLYEVOYRSPFTEWQSKOEN-TCNVITIEGLDAEKCYSEFWVRKAMEDV-YG 195  
 RESULT 10  
 ENTRY  
 TITLE  
 ORGANISM  
 DATE  
 ACCESSIONS  
 REFERENCE  
 #authors  
 #journal  
 #title  
 #cross-references  
 #accession  
 #status  
 #molecule\_type  
 #residues  
 #cross-references  
 #note  
 GENETICS  
 #gene  
 #summary  
 Query Match  
 Best Local Similarity

```

Matches 19; Conservative 18; Mismatches 29; Indels 5; Gaps 5;
Db 4115 NMTNTEIREIKVRAT-D-E-SGIANVTATINGSLSLEKVNETWIGRVELDGKYLENV 4171
Qy 140 DLSYGDLLYEVOYRSPFTEWQSKOEN-TCNVITIEGLDAEKCYSEFWVRKAMEDVYGPDT 198
Db 4172 FASDKWGNVC 4182
Qy 199 YPSD-WSEVTC 208
RESULT 11
ENTRY
TITLE
ALTERNATE_NAMES
CONTAINS
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
#submission
#description
#accession
#molecule_type
#residues
#cross-references
REFERENCE
#authors
#journal
#title
#cross-references
#accession
#status
#molecule_type
#residues
#cross-references
REFERENCE
#authors
#journal
#title
#cross-references
#accession
#status
#molecule_type
#residues
#cross-references
REFERENCE
#authors
#journal
#title
#cross-references
#accession
#status
#molecule_type
#residues
#cross-references
#note
GENETICS
#gene
#map_position
#introns
CLASSIFICATION

```

#type complete  
 twitchin - Caenorhabditis elegans  
 myosin-regulating protein  
 protein kinase (EC 2.7.1.1)  
 #formal\_name Caenorhabditis elegans  
 28-Oct-1995 #sequence\_revision 24-Oct-1997 #text\_change  
 18-Jun-1999  
 S57242; S07571; S06797; S57218  
 S57242  
 Benian, G.M.; L'Hernault, S.W.; Morris, M.E.  
 submitted to the EMBL Data Library, February 1993  
 Additional sequence complexity within twitching of  
 Caenorhabditis elegans muscle.  
 #accession S57242  
 #molecule\_type DNA  
 #residues 1-6839 #label BEN1  
 #cross-references EMBL:L10351  
 #experimental\_source var. Bristol  
 S07571  
 Benian, G.  
 submitted to the EMBL Data Library, November 1989  
 S07571  
 #molecule\_type DNA  
 #residues 792-6839 #label BEN2  
 #cross-references EMBL:X15423; NID:g6897; PID:CAA33463.1; PID:g6898  
 #experimental\_source var. Bristol  
 S06797  
 Benian, G.M.; Kiff, J.E.; Neckelmann, N.; Moerman, D.G.;  
 Waterston, R.H.  
 Nature (1989), 342:45-50  
 Sequence of an unusually large protein implicated in  
 regulation of myosin activity in C. elegans.  
 #cross-references MUID:90044042  
 S06797  
 nucleic acid sequence not shown  
 #molecule\_type DNA  
 #residues 806-1175; 1178-1998, 'Y', 2000-3040, 'I', 3042-3335, 'I',  
 3337-5693; 5696-6359, 'I', 6361-6377; 6386-6478; 6541-6635;  
 6649-6742; 6745-6838 #label BEN3  
 #cross-references EMBL:X15423  
 #experimental\_source var. Bristol  
 S57218  
 Benian, G.M.; L'Hernault, S.W.; Morris, M.E.  
 Genetics (1993) 134:1037-1104  
 Additional sequence complexity in the muscle gene, unc-22,  
 and its encoded protein, twitchin, of Caenorhabditis  
 elegans.  
 #cross-references MUID:93387664  
 S57218  
 #molecule\_type DNA  
 #residues 2-99; 108-194, 'O', 196-206; 374-468; 658-753 #label BEN4  
 #experimental\_source var. Bristol  
 COMMENT Lack of unc-22 leads to a constant twitching of the body muscles.  
 GENETICS  
 #gene unc-22  
 #map\_position IV  
 #introns 18/3; 69/3; 143/2; 176/3; 264/2; 387/3; 413/2; 471/1; 516/3;  
 550/3; 582/3; 603/3; 628/3; 646/3; 674/3; 728/3; 767/2;  
 822/3; 897/1; 1770/1; 2115/3; 2597/2; 2651/1; 2746/1;  
 2820/3; 2948/1; 6152/3; 6691/3; 6776/1; 6808/3  
 #superfamily twitchin; fibronectin type III repeat homology;

immunoglobulin homology; protein kinase homology  
 ARP; autophosphorylation; duplication; muscle;  
 phosphotransferase; serine/threonine-specific pro-  
 tein kinase

## FEATURE

```

FEATURE
806-898,899-990,
991-1083,1084-1175,
1178-1273,
1474-1567,
1770-1864,
2066-2158,
2358-2450,
2651-2745,
2948-3041,
3242-3336,
3536-3629,
3829-3921,
4124-4214,
4517-4611,
4812-4907,
5211-5303,
5304-5398,
5601-5693,
5696-5790,
6263-6356,
6386-6478,
6541-6635,
6649-6742,
6745-6838,
1271-1372,
1373-1473,
1568-1670,
1671-1769,
1865-1964,
1965-2065,
2159-2258,
2259-2357,
2451-2550,
2551-2650,
2746-2847,
2848-2947,
3042-3141,
3142-3241,
3337-3435,
3436-3535,
3630-3729,
3730-3828,
3922-4022,
4023-4123,
4215-4313,
4314-4415,
4416-4516,
4612-4710,
4711-4811,
4908-5009,
5010-5109,
5110-5210,
5399-5499,
5500-5600,
5791-5889,
5940-6197,
5948-5956,
5971
SUMMARY
#region motif 2\
#region motif 1\
#domain protein kinase homology #label KIN\
#region protein kinase Arr-binding motif\
#active-site lys #status predicted
#length 6839 #molecular-weight 753494 #checksum 1785

```

RESULT 12

```

ENTRY          S50579      #type complete
TITLE          hypothetical protein YER076c - yeast (Saccharomyces
                  cerevisiae)
ORGANISM       28-formal_name Saccharomyces cerevisiae
DATE           28-May-1993 #sequence_revision 24-Feb-1995 #text_change
                  21-Nov-1997
ACCESSIONS     S50579
REFERENCE       S50428
                  Dietrich, F.S.
                  #submission submitted to the EMBL Data Library, December 1994
                  #description The sequence of S. cerevisiae cosmids 9747, 8198, 9781, and
                  lambda clones 3612 and 6052.
                  #accession S50579
                  ##molecule_type DNA
                  ##residues 1-302 ##label DIE
GENETICS        ##cross-references EMBL:U18839; NID:g603313; PID:g603314; MIPS:YER076C
                  #map_position 5R
SUMMARY         #length 302 #molecular-weight 33466 #checksum 9755

Query Match          6.3%; Score 97; DB 2; Length 302;
Best Local Similarity 22.8%; Pred. No. 2.03e-01;
Matches 18; Conservative 21; Mismatches 33; Indels 7; Gaps 7;

Db      194  CGSQEFTNIFDQEGWLSFVKTW-ST-NSSCDITASE-GNLTC-AVRYSVSSMHN-HGK 248
           |:  *:::  |:  |:  |:  |:  |:  |:  |:  |:  |:  |:  |:  |:  |:  |:
QY      138  CSDLSYGDLLEYQVY-RSPDFEWSKQDNTCNVTEGLDAEKCYSFVVRKAMEDVYGP 196
           ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

Db      249  TAFCTVYSHGDSW-RAELR 266
           ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
QY      197  DTYPDSWSEVTCWQGEIR 215
           ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

```

## RESULT 13

```

ENTRY      13
TITLE      S77803
ORGANISM   hypothetical protein MG012 - Mycoplasma capricolum (SGC3)
DATE       09-Oct-1997
REFERENCE  #formal_name Mycoplasma capricolum
           #sequence_revision 31-Oct-1997 #text_change
           31-Oct-1997
AUTHORS    S77803; S49950
           S77739
JOURNAL     Bork, P.; Ouzounis, C.; Casari, G.; Schneider, R.; Sander,
           C.; Dolan, M.; Gilbert, W.; Gillevet, P.M.
           Mol. Microbiol. (1995) 16:955-967
TITLE       Exploring the Mycoplasma capricolum genome: a minimal cell
           reveals its physiology.
CROSS-REFS #cross-references MUID:96059641
           #accession S77803
           #status: nucleic acid sequence not shown; translation not shown
           #molecule_type DNA
           #residues 1-311 #label BOR
           #cross-references EMBL:Z33015; NID:959860; PID:g602031
           #experimental_source AFCC 27343
           #note the nucleotide sequence was submitted to the EMBL Data
           Library, July 1994

```

## GENETICS

```

#genetic_code SGC3
SUMMARY      #length 311  #molecular-weight 35768  #checksum 4460

Query Match      6.3%;  Score 97;  DB 2:  Length 311;
Best Local Similarity 27.9%;  Pred. No. 2.03e-01;
Matches 12;  Conservative 11;  Mismatches 19;  Indels 1;  Gaps 1;

Db 202 FCLEVVKATREVIDKVPANFIFGRAPTEPYGDIQYTIED 244
| | | | | : | : | : | : | : | : | : | : | : | : |
QY 37 FNLETVQVWNA-SKYSRTNLIIFYRFNGDEAYDOCTNYLQ 78

```

RESULT 14

ENTRY	TITLE
S12136	#type complete somatotropin receptor precursor - pig



US-09-376-430-2-05.rpr

Thu May 11 06:49:35 2000

ALTERNATE\_NAMES growth hormone receptor  
 ORGANISM #formal\_name Sus scrofa domestica #common\_name domestic pig  
 DATE 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 10-Sep-1997

ACCESSIONS S12136

REFERENCE S12136

#authors Clouff, J.A.; Wang, X.; Kopchick, J.J.

#journal Nucleic Acids Res. (1990) 18:6451

#title Porcine growth hormone receptor cDNA sequence.

#cross-references MUID:91057155

#accession S12136

#status preliminary; translation not shown

#molecule\_type mRNA

#residues 1-638 #label CTO

#cross-references EMBL:X54429; NID:g2037; PID:g2038

KEYWORDS transmembrane protein

SUMMARY #length 638 #molecular-weight 71144 #checksum 2806

Query Match 6.2%; Score 96; DB 2; Length 638;

Best Local Similarity 28.8%; Pred. No. 2.81e-01;

Matches 15; Conservative 15; Mismatches 20; Indels 2; Gaps 2;

Db 190 LEYELQYKVEVNETQKMDPVLSTVPVYSLRLDKXEYVVRVSRQNSSEKYG 241

QY 146 LLYEYQYRSPFDTQWQKQEN-TCNVITIEGLDAEKCYFWVRVKAME-DVYG 195

RESULT 15

ENTRY 156563 #type complete

TITLE interleukin-3 receptor beta-subunit - rat

ORGANISM #formal\_name Rattus sp. #common\_name rat

DATE 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 23-Jul-1999

ACCESSIONS 156563

REFERENCE 156563

#authors Appel, K.; Buttini, M.; Sauter, A.; Gebicke-Haerter, P.J.

#journal J. Neurosci. (1995) 15:5800-5809

#title Cloning of rat interleukin-3 receptor beta-subunit from

cultured microglia and its mRNA expression in vivo.

#cross-references MUID:95370942

#accession 156563

#status preliminary; translated from GB/EMBL/DBJ

#molecule\_type mRNA

#residues 1-896 #label RES

#cross-references GB:S79263; NID:g1086954; PIDN:AAB35068.1;

GENETICS PID:g1086955

#gene rll-3beta

CLASSIFICATION #superfamily interleukin-3 receptor beta chain; cytokine

KEYWORDS receptor homology

cytokine receptor

FEATURE #domain cytokine receptor homology #label CRS1\

39-235 #domain cytokine receptor homology #label CRS2

253-433 #length 896 #molecular-weight 99504 #checksum 5352

SUMMARY

Query Match 6.2%; Score 96; DB 2; Length 896;

Best Local Similarity 30.3%; Pred. No. 2.81e-01;

Matches 20; Conservative 17; Mismatches 21; Indels 8; Gaps 8;

Db 375 FOYQYKKKLDKDR-WEDSKTENLNAHSMDLPLEFGTSTYCARVRVKTIPE-Y-KGLW-SEW 430

QY 148 YEYQYRSPFDTQWQKQEN-TIE-G-LDAEKCYFWVRVKAMEDVYGPDTYPSDW 203

Db 431 SNECTW 436

QY 204 SEVTCW 209

Search completed: Wed May 10 11:44:20 2000

Job time : 20 secs.



\*\*\*\*\*  
 M P S R E H  
 \*\*\*\*\*  
 (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 11:37:25 2000; MasPar time 99.72 Seconds  
 61.995 Million cell updates/sec  
 Tabular output not generated.

Title: >US-09-376-430-2  
 Description: (23-225) from US09376430A.ppt (5 of 25)  
 Perfect Score: 1545  
 Sequence: 1 OGGAAEGVQIIYFNLETV.....VTCWQGEIRDCAETPTTP 203

Scoring table:  
 PAM 150  
 Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%

Database: Listing first 45 summaries  
 swiss-prot38  
 1:swissprot

Statistics: Mean 43.716; Variance 70.603; scale 0.619

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	210	13.6	373	1	CYRG_CANFA	9.03e-24
2	195	12.6	369	1	CYRG_HUMAN	1.44e-20
3	188	12.2	369	1	CYRG_MOUSE	4.28e-19
4	187	12.1	379	1	CYRG_BOVIN	6.93e-19
5	115	7.4	878	1	IL2RG_MOUSE	6.04e-05
6	113	7.3	897	1	CYRG_HUMAN	1.33e-04
7	108	7.0	896	1	CYRG_MOUSE	9.29e-04
8	100	6.5	638	1	GHR_MOUSE	1.86e-02
9	97	6.3	302	1	YE06_YEAST	5.49e-02
10	96	6.2	638	1	GHR_PIG	7.84e-02
11	94	6.1	638	1	GHR_RABIT	1.59e-01
12	92	6.0	361	1	GLNA_PANAR	3.18e-01
13	93	6.0	380	1	IL132_HUMAN	2.25e-01
14	90	5.8	427	1	INTERLEUKIN-13 RECEPTOR	6.28e-01
15	90	5.8	629	1	RIBOSOMAL PROTEIN S6 K	6.28e-01
16	89	5.8	629	1	GHR_BOVIN	8.80e-01
17	89	5.8	1788	1	YP72_CAEEL	8.80e-01
18	88	5.7	562	1	GCVK_HSV6U	1.23e+00
19	88	5.7	638	1	GHR_MACMU	1.23e+00
20	88	5.7	934	1	EAE2_ECOLI	1.23e+00
21	88	5.7	983	1	EPK3_CHICK	1.23e+00
22	87	5.6	418	1	HEML_SALTY	1.71e+00
23	87	5.6	503	1	CYTOCHROME P450 51 (EC	1.71e+00

24	86	5.6	503	1	CP51_PIG	2.37e+00
25	87	5.6	634	1	GHR_SHEEP	1.71e+00
26	87	5.6	918	1	INTERLEUKIN-6 RECEPTOR	1.71e+00
27	87	5.6	939	1	EAE1_ECOLI	3.28e+00
28	85	5.5	458	1	IF3T_FORCA	3.28e+00
29	85	5.5	470	1	GLNA_FREDI	3.28e+00
30	85	5.5	510	1	CD39_MOUSE	3.28e+00
31	85	5.5	902	1	SYGL_YEAST	3.28e+00
32	83	5.4	257	1	VGLG_BRSVL	6.22e+00
33	83	5.4	398	1	FRHA_METVO	6.22e+00
34	83	5.4	503	1	CP51_HUMAN	6.22e+00
35	83	5.4	554	1	ARD1_RAT	6.22e+00
36	83	5.4	574	1	ARD1_HUMAN	6.22e+00
37	84	5.4	1061	1	DPOL_ADE12	6.22e+00
38	83	5.4	1122	1	RPOB_THECE	6.22e+00
39	84	5.4	1447	1	DCC_MOUSE	4.53e+00
40	83	5.4	1447	1	DCC_HUMAN	6.22e+00
41	83	5.4	5255	1	BACA_BACLI	6.22e+00
42	82	5.3	130	1	SM4_MOUSE	8.51e+00
43	82	5.3	371	1	OM25_HAEIN	8.51e+00
44	82	5.3	550	1	PTR2_HUMAN	8.51e+00
45	82	5.3	1584	1	U104_CAEEL	8.51e+00

ALIGNMENTS

RESULT 1  
 ID CYRG\_CANFA STANDARD; PRT: 373 AA.  
 AC P40321.  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C)  
 DE (INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (p64).  
 GN IL2RG.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SPLEEN;  
 RX MEDLINE; 95130114.  
 RA Henthorn P.S.; Somberg R.L.; Fimiani V.M.; Puck J.M.; Patterson D.F.,  
 RA Felsburg P.J.;  
 RT "IL-2R gamma gene microdeletion demonstrates that canine X-linked  
 RT severe combined immunodeficiency is a homologue of the human  
 RT disease";  
 RL Genomics 23:69-74(1994).  
 CC -|- FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF  
 CC INTERLEUKINS.  
 CC -|- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND  
 CC PROBABLY ALSO THE IL-13 RECEPTORS.  
 CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -|- DISEASE: DEFECTS IN IL2RG ARE THE CAUSE OF A CANINE X-LINKED  
 CC SEVERE COMBINED IMMUNODEFICIENCY.  
 CC -|- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -|- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; U04361; AAC48403.1;  
 CC HSPSP; P31785; ILLN.  
 CC PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1  
 CC PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.  
 CC PFAM; PF00041; fn3; 1.  
 CC Receptor; Transmembrane; Glycoprotein; signal.  
 CC SIGNAL 1 22  
 CC POTENTIAL.

FT CHAIN 23 373 CYTOKINE RECEPTOR COMMON GAMMA CHAIN.  
 FT DOMAIN 23 261 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 261 283 POTENTIAL.  
 FT DOMAIN 284 373 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 151 249 FIBRONECTIN TYPE-III.  
 FT DISULFID 62 72 POTENTIAL.  
 FT DISULFID 102 115 POTENTIAL.  
 FT CARBOHYD 24 24 POTENTIAL.  
 FT CARBOHYD 71 71 POTENTIAL.  
 FT CARBOHYD 75 75 POTENTIAL.  
 FT CARBOHYD 84 84 POTENTIAL.  
 FT CARBOHYD 159 159 POTENTIAL.  
 FT CARBOHYD 164 164 POTENTIAL.  
 FT CARBOHYD 249 249 POTENTIAL.  
 SQ SEQUENCE 373 AA; 42516 MW; 03A0DE1FB089D8B CRC64;

Query Match 13.6%; Score 210; DB 1; Length 373;  
 Best Local Similarity 25.9%; Pred. No. 9.03e-24;  
 Matches 49; Conservative 56; Mismatches 73; Indels 11; Gaps 10;

Db 59 EVQCFVFNVEYMCNCTSSSPRPNTLHLHYKNSNDKQVCEGHYLFSEVITAGCWQ 118  
 QY 31 QIQIIFNLEIVQVNTWASKYSR-TNLTFHYRF-NGDEA-YDQCTNYLLQEGHTSGCLLD 87  
 Db 119 KEEHLHYETFFVQLRDPREPRQSTQKLQNLVLPWAPENLTLNLSESOLELSWSNRH 178  
 QY 88 AEQRD--DILYFSIRNGTHPVFTAS-RWNVYL-KPSSPKHVRP-SWHQDAVTVICSDLS 142  
 Db 179 LDHGLEHVQVRSWDSWTSQSDVDRNSFLSPVQDGKFTYFVRYSR-NPLGSAQRW 237  
 QY 143 YGD-LLEVOVRSFPDTEW-OSKQENTCNVTIEGLDAEKYSFWVRKAMEDVVGPDTP 200  
 Db 238 SEWSHPINW 246  
 QY 201 SDWSEVTCW 209

RESULT 2  
 ID CYRG\_HUMAN STANDARD; PRT; 369 AA.  
 AC P31785;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C) (INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64) (CD132 ANTIGEN).  
 GN IL2RG.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE; 92335883.  
 RA Takeshita T., Asao H., Ohtani K., Ishii N., Kumaki S., Tanaka N., Munakata H., Nakamura M., Sugamura K.;  
 RT "Cloning of the gamma chain of the human IL-2 receptor."  
 RL Science 257:379-382(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RX MEDLINE; 93293887.  
 RA Noguchi M., Adelstein S., Cao X., Leonard W.J.;  
 RT "Characterization of the human interleukin-2 receptor gamma chain gene."  
 RL J. Biol. Chem. 268:13601-13608(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND VARIANTS ASP-114 AND ASN-153.  
 RX MEDLINE; 94004847.  
 RA Puck J.M., Deschenes S.M., Porter J.C., Dutra A.S., Brown C.J., Willard H., Henthorn P.S.;  
 RT "The interleukin-2 receptor gamma chain maps to Xq13.1 and is mutated in X-linked severe combined immunodeficiency, SCIDX1.".  
 RL Hum. Mol. Genet. 2:1099-1104(1993).  
 RN [4]

RP IDENTIFICATION AS A IL-4R SUBUNIT.  
 RX MEDLINE; 94090315.  
 RA Kondo M., Takeshita T., Ishii N., Nakamura M., Watanabe S., Arai K.-I., Sugamura K.;  
 RT "Sharing of the interleukin-2 (IL-2) receptor gamma chain between receptors for IL-2 and IL-4.";  
 RL Science 262:1874-1877(1993).  
 RN [5]  
 RP IDENTIFICATION AS A IL-4R SUBUNIT.  
 RX MEDLINE; 94090317.  
 RA Russell S.M., Kkegan A.D., Harada N., Nakamura Y., Noguchi M., Leland P., Friedmann M.C., Miyajima A., Puri R.K., Paul W.E., Leonard W.J.;  
 RT "Interleukin-2 receptor gamma chain: a functional component of the interleukin-4 receptor.";  
 RL Science 262:1880-1883(1993).  
 RN [6]  
 RP IDENTIFICATION AS A IL-7R SUBUNIT.  
 RX MEDLINE; 94090316.  
 RA Noguchi M., Nakamura Y., Russell S.M., Ziegler S.F., Tsang M., Cao X., Leonard W.J.;  
 RT "Interleukin-2 receptor gamma chain: a functional component of the interleukin-7 receptor.";  
 RL Science 262:1877-1880(1993).  
 RN [7]  
 RP 3D-STRUCTURE MODELING OF 57-248.  
 RX MEDLINE; 95111955.  
 RA Bamorough P., Hedgecock C.J., Richards W.G.;  
 RT "The interleukin-2 and interleukin-4 receptors studied by molecular modelling.";  
 RL Structure 2:839-851(1994).  
 RN [8]  
 RP VARIANTS XSCID PHE-115; CYS-240 AND ILE-241.  
 RX MEDLINE; 94130970.  
 RA Disanto J.P., Dautry-Varsat A., Certain S., Fischer A., de Saint Basile G.;  
 RT "Interleukin-2 (IL-2) receptor gamma chain mutations in X-linked severe combined immunodeficiency disease result in the loss of high-affinity IL-2 receptor binding.";  
 RL Eur. J. Immunol. 24:475-479(1994).  
 RN [9]  
 RP VARIANT XSCID LYS-68.  
 RX MEDLINE; 94375038.  
 RA Markiewicz S., Subtil A., Dautry-Varsat A., Fischer A., de Saint Basile G.;  
 RT "Detection of three nonsense mutations and one missense mutation in the interleukin-2 receptor gamma chain gene in SCIDX1 that differently affect the mRNA processing.";  
 RL Genomics 21:291-293(1994).  
 RN [10]  
 RP VARIANT XSCID HIS-162.  
 RX MEDLINE; 94300093.  
 RA Ishii N., Asao H., Kimura Y., Takeshita T., Nakamura M., Tsuchiya S., Konno T., Maeda M., Uchiyama T., Sugamura K.;  
 RT "Impairment of ligand binding and growth signaling of mutant IL-2 receptor gamma-chains in patients with X-linked severe combined immunodeficiency.";  
 RL J. Immunol. 153:1310-1317(1994).  
 RN [11]  
 RP VARIANT XSCID ASN-39.  
 RX MEDLINE; 95023932.  
 RA Disanto J.P., Rieux-Laucat F., Dautry-Varsat A., Fischer A., de Saint Basile G.;  
 RT "Defective human interleukin 2 receptor gamma chain in an atypical X chromosome-linked severe combined immunodeficiency with peripheral T cells.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:9466-9470(1994).  
 RN [12]  
 RP VARIANTS XSCID CYS-226 AND HIS-226.  
 RX MEDLINE; 95397841.  
 RA Pepper A.E., Buckley R.H., Small T.N., Puck J.M.;  
 RT "Two mutational hotspots in the interleukin-2 receptor gamma chain gene causing human X-linked severe combined immunodeficiency.";

Thu May 11 06:49:35 2000

Am. J. Hum. Genet. 57:564-571(1995).  
[13]  
VARIANT XSCID SER-183.  
Clark P.A., Lester T., Genet S., Jones A.M., Hendriks R.,  
Levin R.L., Kinon C.;  
"Screening for mutations causing X-linked severe combined  
immunodeficiency in the IL-2R gamma chain gene by single-strand  
conformation polymorphism analysis";  
Hum. Genet. 96:427-432(1995).  
[14]  
VARIANT XSCID GLN-HIS-TRP INS-237.  
MEDLINE: 95164726.  
Puck J.M., Pepper A.E., Bedard P.-M., Laframboise R.;  
"Female germ line mosaicism as the origin of a unique IL-2 receptor  
gamma-chain mutation causing X-linked severe combined  
immunodeficiency";  
J. Clin. Invest. 95:895-899(1995).  
[15]  
VARIANT XSCID GLN-271.  
MEDLINE: 95190013.  
Schmalstieg F.C., Leonard W.J., Noguichi M., Berg M., Rudloff H.E.,  
Denney R.M., Dave S.K., Brooks E.G., Goldman A.S.;  
"Missense mutation in exon 7 of the common gamma chain gene causes a  
moderate form of X-linked combined immunodeficiency";  
J. Clin. Invest. 95:1169-1173(1995).  
[16]  
VARIANT XSCID ARG-115.  
MEDLINE: 97042245.  
Stephan V., Wahn V., le Deist F., Dirksen U., Broker B.,  
Muller-Fleckenstein I., Horneff G., Schroten H., Fischer A.,  
de Saint Basile G.;  
"Atypical X-linked severe combined immunodeficiency due to possible  
spontaneous reversion of the genetic defect in T cells";  
New Engl. J. Med. 335:1563-1567(1996).  
[17]  
VARIANT XSCID GLN-285.  
MEDLINE: 97295088.  
Jones A.M., Clark P.A., Katz F., Genet S., McMahon C., Alterman L.,  
Cant A., Kinon C.;  
"B-cell-negative severe combined immunodeficiency associated with a  
common gamma chain mutation";  
Hum. Genet. 99:677-680(1997).  
[18]  
VARIANT XSCID CYS-222.  
MEDLINE: 98064061.  
Sharif N., Shahar M., Roifman C.M.;  
"An interleukin-2 receptor gamma chain mutation with normal thymus  
morphology";  
J. Clin. Invest. 100:3036-3043(1997).  
CC -1- FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF  
INTERLEUKINS.  
CC -1- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND  
PROBABLY ALSO THE IL-13 RECEPTORS.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- DISEASE: DEFECTS IN IL2RG ARE THE CAUSE OF A SEVERE COMBINED  
IMMUNODEFICIENCY, WHICH IS KNOWN AS AGAMAGLOBULINEMIA, SWISS TYPE  
OR X-LINKED SEVERE COMBINED IMMUNODEFICIENCY DISEASE (XSCID).  
CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
CC -1- DATABASE: NAME-PROW; NOTE-CD guide CD132 entry;  
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd132.htm".  
CC -1- DATABASE: NAME-IL2RGbase; NOTE-X-linked SCID mutation database;  
WWW="http://www.nhgri.nih.gov/DIR/LGT/SCID/IL2RGbase.html".  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL: D11086; BAA01857.1; -.  
EMBL: L12183; AAA59145.1; -.  
EMBL: L12178; AAA59145.1; JOINED.  
EMBL: L12176; AAA59145.1; JOINED.  
EMBL: L12177; AAA59145.1; JOINED.  
EMBL: L12179; AAA59145.1; JOINED.  
EMBL: L12180; AAA59145.1; JOINED.  
EMBL: L12181; AAA59145.1; JOINED.  
EMBL: L12182; AAA59145.1; JOINED.  
EMBL: L19546; AAC37524.1; -.  
PDB: A42565; A42565.  
PDB: 1ILM; 26-JAN-95.  
PDB: 1ILN; 26-JAN-95.  
PDB: 300400; -.  
PDB: 308980; -.  
PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
PROSITE; PS00440; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.  
PFAM; PF00041; fn3; 1.  
KW Receptor; Transmembrane; Glycoprotein; Signal; Disease mutation;  
...  
Note: remainder of annotations omitted.  
Query Match 12.6%; Score 195; DB 1; Length 369;  
Best Local Similarity 25.9%; Pred. No. 1.44e-20;  
Matches 49; Conservative 52; Mismatches 77; Indels 11; Gaps 11;  
Db 59 EVQCFVFNVEYMNCTNSSESPQPTNLTHYWKNSDNDKVKCHYLFSEITSCGLQ 118  
QY 31 QIQIIVFNLETQVTVNASKYSR-TNLTPHYRF-NGD-EAYDOCTNYLQEGHTSGCLLD 87  
Db 119 KEIHLHYQFVVOLODPRPRQATQMLKQLNLPWAPENLTLLKLSQLELNNRNF 178  
QY 88 AEQRDDILYFSIR-NGTH-PVETASRW-VYVL-KPSSPKHVRF-SWHQDAVITVTCSDL 142  
Db 179 LNHCEHLHYQYRTDWDHWSVEOSDYRHRKFLSPVDGQKRYTRFVRSR-FNPLCGSAQHW 237  
QY 143 YGD-LLYEVOYRSPFDEW-QSKQENTCNVTIEGLDAEKCYSFWRVVKMEDVYGDTP 200  
Db 238 SEWSHPFIHW 246  
QY 201 SDWSEVTCW 209  
RESULT 3  
ID CYRG\_MOUSE STANDARD; PRT; 369 AA.  
AC P34902;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C)  
DE (INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64).  
GN IL2RG.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 93377575.  
RA Kumaki S., Kondo M., Takeshita T., Asao H., Nakamura M., Sugamura K.;  
RT "Cloning of the mouse interleukin 2 receptor gamma chain:  
RT demonstration of functional differences between the mouse and human  
RT receptors";  
RL Biochem. Biophys. Res. Commun. 193:356-363(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CH/CA;  
RX MEDLINE: 93391374.  
RA Cao X., Kozak C.A., Liu Y.J., Noguichi M., O'Connell F., Leonard W.J.;  
RT "Characterization of cDNAs encoding the murine interleukin 2 receptor  
RT (IL-2R) gamma chain: chromosomal mapping and tissue specificity of  
RT IL-2R gamma chain expression";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:8464-8468(1993).  
RN [3]

```

FT  CARBOHYD  164  164  POTENTIAL.
SQ  SEQUENCE  369 AA;  42241 MW;  CB2D5AB459077AC7 CRC64;

Query Match      12.2%;  Score 188;  DB 1;  Length 369;
Best Local Similarity 26.8%;  Pred. NO. 4.28e-19;
Matches      51;  Conservative      52;  Mismatches 75;  Indels 12;  Gaps 11;

Db  59  EVQCFVFNIEYMCNTWNSSEPAQNLTHRYRYKYSDNNNTFOECSHYLSFKETSGCOIQ 118
    ::  |||  ::  |||  ::  |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
QY  31  QIQIIVFNLETQVTTWNASKYSR-INLFIHYEFN-GDE-AYDOCTNYLLQEGHTSGCLLD 87
    ::  |||  ::  |||  ::  |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

Db  119  KEDIQLYQTFVVLQDPQPKQPRRAVQKLNQLNLVTPRAPENLTLSNLSESOLELRWKSRR 178
    ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
QY  88  AKQRDDILYFSIR-NGTH-PVFTASRWV-VYYL-KPSSPKHVRES-WHQDAVTV--TCSD 140
    ::  |||  ::  |||  ::  |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

Db  179  IKERCILQYLQVQYRSNRDRSWTELIIVNHEPRSLPSVDLKYTRFVRSRY-NPTICGSSQ 237
    ::  |||  ::  |||  ::  |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
QY  141  LSYGDLLEYQYRSFDFEW-QSKQENTCVNTIEGLDAEKYSEFWRVKAMEDVYGPDTY 199
    ::  |||  ::  |||  ::  |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

Db  238  WSKWSQPVHW 247
    ::  |||  ::  |||  ::  |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
QY  200  PSDNSEVTCW 209
    ::  |||  ::  |||  ::  |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::

RESULT 4
ID  CYRG BOVIN  SEQUENCE  STANDARD;  PRT;  379 AA.
AC  Q95118;
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DI  01-NOV-1997 (Rel. 35, Last sequence update)
DE  15-JUL-1998 (Rel. 36, Last annotation update)
DE  CYTOKINE RECEPTOR COMMON GAMMA CHAIN PRECURSOR (GAMMA-C)
DE  (INTERLEUKIN-2 RECEPTOR GAMMA CHAIN) (IL-2R GAMMA CHAIN) (P64).
GN  IL2RG
OS  Bos taurus (Bovine).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC  Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC  Bovinae; Bos.
RN  [1]
RS  SEQUENCE FROM N.A.
RX  MEDLINE: 96268473.
RA  Yoo J., Stone R.T., Solinas-Toldo S., Fries R., Beattie C.W.;
RA  "Cloning and chromosomal mapping of bovine interleukin-2 receptor
RA  gamma gene.";
RA  DNA Cell Biol. 15:453-459(1996).
CC  -!- FUNCTION: COMMON SUBUNIT FOR THE RECEPTORS FOR A VARIETY OF
CC  INTERLEUKINS.
CC  -!- SUBUNIT: THE GAMMA CHAIN IS COMMON TO THE IL-2, IL-4, IL-7 AND
CC  PROBABLY ALSO THE IL-13 RECEPTORS.
CC  -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC  -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
CC  -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: U33748; AAB07812.1;
CC  HSP: P31785; IILN..
CC  PROSITE: PS00341; RECEPTOR_CYTOKINES_1; 1.
CC  PROSITE: PS00340; RECEPTOR_CYTOKINES_2; FALSE_NEG.
CC  PFAM: PF00041; fn3; 1.
CC  Receptor; Transmembrane; Glycoprotein; Signal.
CC  SIGNAL 1 22
CC  T CHAIN 23 379
CC  T DOMAIN 23 269
CC  T TRANSMEM 270 290
CC  T DOMAIN 291 379
CC  T DOMAIN 158 256
CC  T DISULFID 68 78
CC  POTENTIAL.
CC  CYTOKINE RECEPTOR COMMON GAMMA CHAIN.
CC  EXTRACELLULAR (POTENTIAL).
CC  POTENTIAL.
CC  CYTOPLASMIC (POTENTIAL).
CC  FIBRONECTIN TYPE-III.
CC  POTENTIAL.

```

US-09-376-430-2-05.isp

Thu May 11 06:49:35 2000

```
FT DISULFID 109 122 POTENTIAL.
FT CARBOHYD 77 77 POTENTIAL.
FT CARBOHYD 81 81 POTENTIAL.
FT CARBOHYD 90 90 POTENTIAL.
FT CARBOHYD 166 166 POTENTIAL.
FT CARBOHYD 171 171 POTENTIAL.
SQ SEQUENCE 379 AA; 43037 MW; 33CFAD9C9B032178 CRC64;

Query Match 12.1%; Score 187; DB 1; Length 379;
Best Local Similarity 29.3%; Pred. No. 6.93e-19;
Matches 61; Conservative 49; Mismatches 79; Indels 19; Gaps 17;

Db 66 VOCFVFNVEYVNCWNSSEPPQNNLTHYGRNFNGDKLQEGCHLFSFGITSGCW- 124
QY 32 IQIIFYNLETQVTVNASKYSKTN-LTFHY--R-FNGDEYDOCTNYLLQEGHTSGCLLD 87

Db 125 G-KKEIRLYETFFVOLQDPREHRKQPKMLQDLVWPAPENLTIRNLSEFQLELSWSN 183
QY 88 AEQRDDILY--FSIR-NGTHPVFTASRWV-YY-L-KPSPKHVRFSWHQD-AVTVTGSD 140

Db 184 -RYDKHLEHLVYVESDRDRSQTQSDVDRHSFSLPSVDAOKLYTFVRSRY-NPLCGSA 241
QY 141 LSYGDL-L-YEVTRSPDTEW-QSKQENTCNV-TIE-G-LDAKCYSFVVRKAMEDVYGP 197

Db 242 QHSDMSYPIHWGNSNTSKEN-IENPENP 268
QY 198 TYPDSEVTCWQGEIRDCAETPTTP 225

RESULT 5
ID IL3B_MOUSE STANDARD; PRT; 878 AA.
AC P26954;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE INTERLEUKIN-3 RECEPTOR CLASS II BETA CHAIN PRECURSOR (COLONY
DE STIMULATING FACTOR 2 RECEPTOR, BETA 2 CHAIN).
GN CSF2RB2 OR A12CA OR IL3RB2 OR IL3R.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90117145.
RA Yahara I., Arai K., Miyajima A., Gorman D.M., Maruyama K., Ishii A.,
RA "Cloning of an interleukin-3 receptor gene: a member of a distinct
RT receptor gene family.";
RL Science 247:324-327(1990).
CC -!- FUNCTION: IN MOUSE THERE ARE TWO CLASSES OF HIGH-AFFINITY IL-3
CC RECEPTORS. ONE CONTAINS THIS IL-3-SPECIFIC BETA CHAIN AND THE
CC OTHER CONTAINS THE BETA CHAIN ALSO SHARED BY HIGH-AFFINITY IL-5
CC AND GM-CSF RECEPTORS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; M29855; AAA39295.1;
CC DR PIR; A40091; A40091.
CC DR MGD; MG11339760; CSF2RB2.
CC DR PROSITE; PS00241; RECEPTOR_CYTOKINES_1; 1.
CC DR PRAM; PF00041; fn3; 2.
CC DR KW Receptor; Transmembrane; Glycoprotein; Signal.
CC DR SIGNAL 1 22 POTENTIAL.
```

```
FT CHAIN 23 878 INTERLEUKIN-3 RECEPTOR CLASS II BETA
FT DOMAIN 23 440 CHAIN.
FT TRANSMEM 441 462 EXTRACELLULAR (POTENTIAL).
FT DISULFID 463 499 POTENTIAL.
FT DISULFID 39 49 BY SIMILARITY.
FT DISULFID 78 95 BY SIMILARITY.
FT DISULFID 254 264 BY SIMILARITY.
FT DISULFID 293 310 BY SIMILARITY.
FT CARBOHYD 62 62 POTENTIAL.
FT CARBOHYD 350 350 POTENTIAL.
SQ SEQUENCE 878 AA; 97195 MW; 8EBC9092ADC24D56 CRC64;

Query Match 7.4%; Score 115; DB 1; Length 878;
Best Local Similarity 32.4%; Pred. No. 6.04e-05;
Matches 23; Conservative 17; Mismatches 23; Indels 8; Gaps 8;

Db 370 YIDHTFQVQYKKKSES-WKDSKTKENLGRVNSMDLPQLEPDTSYCARVVRKESD-YD-GI 426
QY 143 YGDLLEYQVYRSPDTEW-QSKQENTCNV-TIE-G-LDAKCYSFVVRKAMEDVYGP 198

Db 427 W-SEWSNEYTW 436
QY 199 YPSDWSEVTCW 209

RESULT 6
ID CYRB_HUMAN STANDARD; PRT; 897 AA.
AC P32927;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CYTOKINE RECEPTOR COMMON BETA CHAIN PRECURSOR (CD131 ANTIGEN).
GN CSF2RB OR IL3RB OR IL3RB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91088571.
RA Hayashida K., Kitamura T., Gorman D.M., Arai K., Yokota T.,
RA Miyajima A.;
RT "Molecular cloning of a second subunit of the receptor for human
RT granulocyte-macrophage colony-stimulating factor (GM-CSF):
RT reconstitution of a high-affinity GM-CSF receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:9655-9659(1990).
RN [2]
RP REVISION TO 454.
RA Kitamura T.;
RA Submitted (FEB-1991) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HIGH AFFINITY RECEPTOR FOR INTERLEUKIN-3, INTERLEUKIN-5
CC AND GRANULOCYTE-MACROPHAGE COLONY-STIMULATING FACTOR.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. THE BETA
CC CHAIN IS COMMON TO THE IL-3, IL-5 AND GM-CSF RECEPTORS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.
CC -!- DATABASE: NAME-PROT; NOTE-CD guide CDw131 entry;
CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cdw131.htm
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; M59941; AAA18171.1;
CC DR PIR; A39255; A39255.
CC DR MIM; 136981.
CC DR PROSITE; PS00241; RECEPTOR_CYTOKINES_1; 2.
CC DR PROSITE; PS00340; RECEPTOR_CYTOKINES_2; 1.
```

DR PFAM: PF00041; fn3; 2.  
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.  
 FT SIGNAL 1 16  
 FT CHAIN 17 897  
 FT DOMAIN 17 443  
 FT TRANSMEM 444 460  
 FT DOMAIN 461 897  
 FT DOMAIN 129 238  
 FT DOMAIN 336 434  
 FT DISULFID 35 45  
 FT CARBOHYD 75 91  
 FT CARBOHYD 58 58  
 FT CARBOHYD 191 191  
 FT CARBOHYD 346 346  
 SQ SEQUENCE 897 AA; 97335 MW; 3398E37FDB8F393A CRC64;

Query Match 7.3%; Score 113; DB 1; Length 897;  
 Best Local Similarity 31.9%; Pred. No. 1.33e-04;  
 Matches 22; Conservative 10; Mismatches 34; Indels 3; Gaps 3;

Db 372 FEIQYKDAT-WKDEKTEFLQNAHSMALPALEPSTFYWARVVRTSGTNGIWSWSE 430  
 QY 148 YEQYRSPDTEW-QSKQENTCNVTIEGLDA-EKCYSEFVVRKAMEDVIGPDYPSDSE 205  
 Db 431 ARSWDTESV 439  
 QY 206 VTCWQGEI 214

RESULT 7  
 ID CYRB\_MOUSE STANDARD; PRT; 896 AA.  
 AC P26955;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-FEB-2000 (Rel. 39, Last annotation update)  
 DE CYTOKINE RECEPTOR COMMON BETA CHAIN PRECURSOR.  
 GN CSF2RB OR CSF2B1 OR AIC2B OR IL3RB1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90319131.  
 RA Gorman D.M., Itoh N., Kitamura T., Schreurs J., Yonehara S.,  
 RA Yahara I., Arai K., Miyajima A.;  
 RT "Cloning and expression of a gene encoding an interleukin 3 receptor-  
 RT like protein: identification of another member of the cytokine  
 RT receptor gene family.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5459-5463(1990).  
 CC -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR INTERLEUKIN-3, INTERLEUKIN-5  
 CC AND GRANULOCYTE-MACROPHAGE COLONY-STIMULATING FACTOR.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. THE BETA  
 CC CHAIN IS COMMON TO THE IL-3, IL-5 AND GM-CSF RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M34397; AAA37204.1;  
 DR PIR: A35782; A35782.  
 DR MGD; MGI:1339759; CSF2RB1.  
 DR PROSITE; PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
 DR PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; 1.  
 DR PFAM; PF00041; fn3; 2.  
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.  
 FT SIGNAL 1 22  
 FT POTENTIAL.

FT CHAIN 23 896  
 FT DOMAIN 23 441  
 FT TRANSMEM 442 463  
 FT DOMAIN 464 896  
 FT DOMAIN 132 241  
 FT DOMAIN 343 440  
 FT DISULFID 39 49  
 FT DISULFID 77 94  
 FT CARBOHYD 62 62  
 FT CARBOHYD 141 141  
 FT CARBOHYD 350 350  
 SQ SEQUENCE 896 AA; 99111 MW; 8CE16EDFDC07A999 CRC64;

Query Match 7.0%; Score 108; DB 1; Length 896;  
 Best Local Similarity 30.6%; Pred. No. 9.29e-04;  
 Matches 22; Conservative 18; Mismatches 24; Indels 8; Gaps 8;

Db 370 SFIEHTFOVYKKKSDS-WEDSKTENLDRAHSMDSQLEPDTSYCARVRVKPISN-YD-G 426  
 QY 142 SYGDLLEYEQYRSPDTEWQ-SKQEN-TCNVITIE-G-LDAEKCYSEFVVRKAMEDVYGP 197  
 Db 427 IW-SKWSEYTW 437  
 QY 198 TYPDWDSEVTCW 209

RESULT 8  
 ID GHR\_HUMAN STANDARD; PRT; 638 AA.  
 AC P10912;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE GROWTH HORMONE RECEPTOR PRECURSOR (GH RECEPTOR) (SERUM BINDING  
 DE PROTEIN).  
 GN GHR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=LIVER;  
 RX MEDLINE; 88065896.  
 RA Leung D.W., Spencer S.A., Cachianes G., Hammonds R.G., Collins C.,  
 RA Henzel W.J., Barnard R., Waters M.J., Wood W.I.;  
 RT "Growth hormone receptor and serum binding protein: purification,  
 RT cloning and expression.";  
 RL Nature 330:537-543(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90046742.  
 RA Godowski P.J., Leung D.W., Meacham L.R., Galgani J.P., Hellmiss R.,  
 RA Keret R., Rotwein P.S., Parks J.S., Larson Z., Wood W.I.;  
 RT "Characterization of the human growth hormone receptor gene and  
 RT demonstration of a partial gene deletion in two patients with Laron-  
 RT type dwarfism.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:8083-8087(1989).  
 RN [3]  
 RP DISULFIDE BONDS.  
 RX MEDLINE; 90153957.  
 RA Fuh G., Mulkerin M.G., Bass S., McFarland N., Brochier M.,  
 RA Bourrel J.H., Light D.R., Wells J.A.;  
 RT "The human growth hormone receptor. Secretion from Escherichia coli  
 RT and disulfide bonding pattern of the extracellular binding domain.";  
 RL J. Biol. Chem. 265:3111-3115(1990).  
 RN [4]  
 RP VARIANT-LARON DWARFISM SER-114.  
 RX MEDLINE; 89384829.  
 RA Amselem S., Duquesnoy P., Attree O., Novelli G., Bousnina S.,  
 RA Postelvinay M.-C., Goossens M.;  
 RT "Laron dwarfism and mutations of the growth hormone-receptor gene.";  
 RL New Engl. J. Med. 321:989-995(1989).  
 RN [5]  
 RP VARIANTS LARON DWARFISM.



Thu May 11 06:49:35 2000

RX MEDLINE: 93278381.  
 RA Anselm S., Duquesnoy P., Duriez B., Dastot F., Sorbier M.-L.,  
 RA Vallet S., Goossens M.;  
 RA "Spectrum of growth hormone receptor mutations and associated  
 RT haplotypes in Laron syndrome.";  
 RL Hum. Mol. Genet. 2:355-359(1993).  
 RP [6]  
 RP VARIANT LARON DWARFISM HIS-170.  
 RX MEDLINE: 94185645.  
 RA Duquesnoy P., Sorbier M.-L., Duriez B., Dastot F., Buchanan C.R.,  
 RA Savage M.O., Preece M.A., Craescu C.T., Blouquit Y., Goossens M.,  
 RA Anselm S.;  
 RA "A single amino acid substitution in the extracellular domain of the  
 RT human growth hormone (GH) receptor confers familial GH resistance  
 RT (Laron syndrome) with positive GH-binding activity by abolishing  
 RT receptor homodimerization.";  
 RL EMBO J. 13:1386-1395(1994).  
 RP [7]  
 RP VARIANTS IDIOPATHIC SHORT STATURE LYS-62; CYS-179 AND ASP-242.  
 RX MEDLINE: 96013502.  
 RA Goddard A.D., Covello R., Luch S.-M., Clarkson T., Attie K.M.,  
 RA Gundheit N., Rundle A.C., Wells J.A., Carlsson L.M.S.;  
 RA "Mutations of the growth hormone receptor in children with idiopathic  
 RT short stature. The Growth Hormone Insensitivity Study Group.";  
 RL New Engl. J. Med. 333:1093-1098(1995).  
 RP [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 19-256.  
 RX MEDLINE: 92196577.  
 RA de Vos A.M., Ullrich M., Kossiakoff A.A.;  
 RA "Human growth hormone and extracellular domain of its receptor:  
 RT crystal structure of the complex.";  
 RL Science 255:306-312(1992).  
 RP [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 19-256.  
 RX MEDLINE: 97113023.  
 RA Sundstrom M., Lundqvist T., Rodin J., Giebel L.B., Milligan D.,  
 RA Norstedt G.;  
 RA "Crystal structure of an antagonist mutant of human growth hormone,  
 RT G120R, in complex with its receptor at 2.9-A resolution.";  
 RL J. Biol. Chem. 271:32197-32203(1996).  
 CC -!- FUNCTION: THIS IS A RECEPTOR FOR PITUITARY GLAND GROWTH HORMONE.  
 CC -!- SUBUNIT: HOMODIMER.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- DISEASE: DEFICIENCY IN GHR IS THE CAUSE OF PITUITARY DWARFISM II  
 CC -!- (LARON-TYPE PITUITARY DWARFISM OR LARON SYNDROME (LS)). IT ALSO  
 CC CAUSES IDIOPATHIC SHORT STATURE.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X06562; CAA29808.1;  
 DR EMBL: M28466; AAA52555.1;  
 DR EMBL: M28458; AAA52555.1; JOINED.  
 DR EMBL: M28459; AAA52555.1; JOINED.  
 DR EMBL: M28460; AAA52555.1; JOINED.  
 DR EMBL: M28461; AAA52555.1; JOINED.  
 DR EMBL: M28462; AAA52555.1; JOINED.  
 DR EMBL: M28463; AAA52555.1; JOINED.  
 DR EMBL: M28464; AAA52555.1; JOINED.  
 DR EMBL: M28465; AAA52555.1; JOINED.  
 DR PIR: S04530; S04530.  
 DR PIR: A33991; A33991.  
 DR PDB: 3HHR; 30-APR-94.  
 DR PDB: 1HWG; 19-NOV-97.  
 DR PDB: 1LHW; 19-NOV-97.  
 DR PDB: 1AXI; 28-JAN-98.

DR PDB: 1A22; 29-APR-98.  
 DR MIM: 600946; -.  
 DR MIM: 262500; -.  
 DR PROSITE: PS00441; RECEPTOR\_CYTOKINES\_1; 1.  
 DR PROSITE: PS00340; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.  
 DR PFAM: PF00041; fn3; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Signal; 3D-structure;  
 KW Dwarfism; Disease mutation.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 638 GROWTH HORMONE RECEPTOR.  
 FT DOMAIN 19 264 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 265 288 POTENTIAL.  
 FT DOMAIN 289 638 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 143 252 FIBRONECTIN TYPE-III.  
 FT DISULFID 56 66  
 FT DISULFID 101 112  
 FT DISULFID 126 140  
 FT CARBOHYD 46 46  
 FT CARBOHYD 115 115  
 FT CARBOHYD 156 156  
 FT CARBOHYD 161 161  
 FT CARBOHYD 200 200  
 FT VARIANT 62 62  
 FT VARIANT 89 89 E -> K (IN IDIOPATHIC SHORT STATURE).  
 FT VARIANT 114 114 R -> K (IN LARON-DWARFISM).  
 FT VARIANT 143 143 F -> S (IN LARON-DWARFISM).  
 FT VARIANT 162 162 V -> A (IN LARON-DWARFISM).  
 FT VARIANT 170 170 V -> D (IN LARON-DWARFISM).  
 FT VARIANT 179 170 D -> H (IN LARON-DWARFISM; ABOLISH  
 FT RECEPTOR HOMODIMERIZATION).  
 FT VARIANT 179 179 R -> C (IN LARON-DWARFISM, AND IDIOPATHIC  
 FT SHORT STATURE).  
 FT VARIANT 229 229 R -> G (IN LARON-DWARFISM).  
 FT VARIANT 242 242 E -> D (IN IDIOPATHIC SHORT STATURE).  
 FT CONFLICT 544 544  
 FT STRAND 53 58  
 FT STRAND 64 68  
 FT STRAND 82 88  
 FT STRAND 99 100  
 FT TURN 104 107  
 FT TURN 109 110  
 FT STRAND 111 114  
 FT TURN 116 117  
 FT STRAND 124 131  
 FT TURN 132 133  
 FT STRAND 134 142  
 FT STRAND 143 145  
 FT HELIX 147 147  
 FT STRAND 153 162  
 FT TURN 164 165  
 FT STRAND 168 176  
 FT TURN 179 180  
 FT TURN 183 186  
 FT STRAND 190 198  
 FT TURN 199 200  
 FT STRAND 205 206  
 FT STRAND 210 210  
 FT STRAND 214 221  
 FT TURN 222 223  
 FT STRAND 225 234  
 FT TURN 247 250  
 SQ SEQUENCE 538 AA; 71499 MW; EAF77EAD4787822 CRC64;

Query Match 6.5%; Score 100; DB 1; Length 638;  
 Best Local Similarity 32.7%; Pred. No. 1.86e-02;

Matches 17; Conservative 11; Mismatches 22; Indels 2; Gaps 2;

Db 190 LEVELOKVEVNETKMKMDPILTSVPVSLKVDKEVVRVRSKORNSGNYG 241  
 QY 146 LLYEVOYRSPDTEWQSKQEN-TCNVITIEGLDAEKCYSFWRVVRKAMEDV-YG 195

RESULT 9  
 ID YE06\_YEAST STANDARD; PRT; 302 AA.  
 AC P40049;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HYPOTHETICAL 33.5 KD PROTEIN IN PTP3-1LV1 INTERGENIC REGION  
 DE PRECURSOR.  
 GN YER076C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;  
 OC Saccharomycetaceae; Saccharomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;  
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
 RA Chung E., Duncan M., Gisman E., Hartzell G., Hunnicke-Smith S.,  
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew D.,  
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,  
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.,  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: TO YEAST KILLER TOXIN KIR.  
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U18839; RA64631.1; --  
 DR Hypothetical protein; Glycoprotein; Signal.  
 FT SIGNAL 1 302  
 FT CHAIN 24 302  
 FT CARBOHYD 65 65  
 FT CARBOHYD 86 86  
 FT CARBOHYD 93 93  
 FT CARBOHYD 220 220  
 FT CARBOHYD 231 231  
 SQ SEQUENCE 302 AA; 33466 MW; 43E3AC5FA2BF378F CRC64;

Query Match 6.3%; Score 97; DB 1; Length 302;  
 Best Local Similarity 22.8%; Pred. No. 5.49e-02;  
 Matches 18; Conservative 21; Mismatches 33; Indels 7; Gaps 7;

Db 194 CGSQEETNFFDQEGWSLFWKW-ST-NSCSDITASE-GNLTC-AVRVSVSSMHN-HGK 248  
 QY 138 CSDLSTGLLYEVOY-RSPDTEWQSKQENTCNVTIEGLDAEKCYSFWRVVRKAMEDVYGP 196  
 Db 249 TAFCVTYSRSDSW-RAELR 266  
 QY 197 DTVPDSDSEVTCQGEIR 215

RESULT 10  
 ID GHR\_PIG STANDARD; PRT; 638 AA.  
 AC P19756;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE GROWTH HORMONE RECEPTOR PRECURSOR (GH RECEPTOR) (SERUM BINDING PROTEIN).  
 DE GHR.

OS Sus scrofa (pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LANDRACE-YORKSHIRE; TISSUE-LIVER;  
 RA MEDLINE; 91057155.  
 RA Clodfi J.A., Wang X., Kopchick J.J.;  
 FT "Porcine growth hormone receptor cDNA sequence.";  
 RL Nucleic Acids Res. 18:6451-6451(1990).  
 CC -!- FUNCTION: THIS IS A RECEPTOR FOR PITUITARY GLAND GROWTH HORMONE.  
 CC -!- SUBUNIT: HOMODIMER.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X54429; CAA38301.1; --  
 DR PIR: S12136; S12136.  
 DR HSSP: P10912; IAZ2.  
 DR PROSITE: PS00340; RECEPTOR\_CYTOKINES\_2; FALSE\_NEG.  
 DR PROSITE: PS00241; RECEPTOR\_CYTOKINES\_1; 1.  
 DR PFAM: PF00041; fn3; 1.  
 DR Receptor; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 638  
 FT DOMAIN 19 284  
 FT TRANSMEM 285 288  
 FT DOMAIN 289 638  
 FT DOMAIN 145 252  
 FT DISULFID 56 66  
 FT DISULFID 104 112  
 FT DISULFID 126 140  
 FT CARBOHYD 46 46  
 FT CARBOHYD 115 115  
 FT CARBOHYD 156 156  
 FT CARBOHYD 161 161  
 FT CARBOHYD 200 200  
 SQ SEQUENCE 638 AA; 71145 MW; BC7C66536F4DF97 CRC64;

Query Match 6.2%; Score 96; DB 1; Length 638;  
 Best Local Similarity 28.8%; Pred. No. 7.84e-02;  
 Matches 15; Conservative 15; Mismatches 20; Indels 2; Gaps 2;

Db 190 LEVELOKVEVNETKMKMDPVLSTVPVSLKVDKEVVRVRSKORNSGNYG 241  
 QY 146 LLYEVOYRSPDTEWQSKQEN-TCNVITIEGLDAEKCYSFWRVVRKAMEDVYGP 195

RESULT 11  
 ID GHR\_RABIT STANDARD; PRT; 638 AA.  
 AC P19941;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE GROWTH HORMONE RECEPTOR PRECURSOR (GH RECEPTOR) (SERUM BINDING PROTEIN).  
 DE GHR.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE; 88065896.  
 RA Leung D.W., Spencer S.A., Cachianes G., Hammonds R.G., Collins C.,  
 RA Henzel W.J., Barnard R., Waters M.J., Wood W.I.;

[illegible]

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X95302; CAA64617.1; -;  
 DR EMBL; U70981; AAB17170.1; -;  
 DR EMBL; Y08768; CAA70021.1; -;  
 DR MIM; 300130; -;  
 DR PROSITE; PS00340; RECEPTOR\_CYTOKINES\_2; 1.  
 DR PFAM; PF00041; fn3; 1.  
 KW Receptor; transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 26  
 FT CHAIN 27 380  
 FT DOMAIN 27 343  
 FT TRANSMEM 344 363  
 FT DOMAIN 364 380  
 FT DISULFID 145 155  
 FT DISULFID 184 197  
 FT CARBOHYD 115 115  
 FT CARBOHYD 215 215  
 FT CARBOHYD 290 290  
 FT CARBOHYD 299 299  
 SQ SEQUENCE 380 AA; 44176 MW; 3C6ACBIB5562C887 CRC64;

Query Match 6.0%; Score 93; DB 1; Length 380;  
 Best Local Similarity 28.8%; Pred. No. 2.25e-01;  
 Matches 19; Conservative 19; Mismatches 23; Indels 5; Gaps 4;  
 Db 133 QGIPETKQVDMCVYXWQYLCSKPKGVLDDNINLFYWEGLDHAL-QCVDYIKAD 191  
 QY 23 QGGAEGVQ-IQIIYFNLETQVQTNWASK--YSRTNLTFFHYRFG-DEAYDQCTNVLQEQ 78  
 Db 192 GONIGC 197  
 QY 79 GHTSGC 84

RESULT 14  
 ID IL131\_HUMAN STANDARD; PRT; 427 AA.  
 AC P78552; Q99656;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE INTERLEUKIN-13 RECEPTOR ALPHA-1 CHAIN PRECURSOR (IL-13R-ALPHA-1) (IL-13RA-1).  
 GN IL13RA1 OR IL13RA OR IL13R.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-CARCINOMA;  
 RX MEDLINE; 97165986.  
 RA Miloux B., Laurent P., Bonnin O., Lupker J., Caput D., Vita N.,  
 RA Ferrara P.;  
 RT "Cloning of the human IL-13R alpha chain and reconstitution with the  
 RT IL4R alpha of a functional IL-4/IL-13 receptor complex";  
 RL FEBS Lett. 401:163-166(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-B-CELL;  
 RA Gauchat J.F.M., Schlagenhauf E., Feng N.P., Moser R., Yamase M.,  
 RA Jeannin P., Alouani S., Elson G., Notarangelo L.D., Wells T.,  
 RA Eugster H.P., Bonnefoy J.Y.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-T-CELL;

RX MEDLINE; 97067184.  
 RA Anan M.J., Tayebi N., Obiri N.I., Puri R.K., Modi W.S.,  
 RA Leonard W.J.;  
 RT "CDNA cloning and characterization of the human interleukin 13  
 RT receptor alpha chain";  
 RL J. Biol. Chem. 271:29265-29270(1996).  
 CC -!- FUNCTION: BINDS IL-13 WITH A LOW AFFINITY. TOGETHER WITH IL-4R-  
 CC ALPHA CAN FORM A FUNCTIONAL RECEPTOR FOR IL-13. ALSO SERVES AS AN  
 CC ALTERNATE ACCESSORY PROTEIN TO THE COMMON CYTOKINE RECEPTOR GAMMA  
 CC CHAIN FOR IL-4 SIGNALING, BUT CANNOT REPLACE THE FUNCTION OF  
 CC GAMMA-C IN ALLOWING ENHANCED IL-2 BINDING ACTIVITY.  
 CC -!- SUBUNIT: INTERLEUKIN-13 RECEPTOR IS A COMPLEX OF IL4R-ALPHA,  
 CC IL13R-ALPHA, AND POSSIBLY OTHER COMPONENTS.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- TISSUE SPECIFICITY: UBIQUITOUS. HIGHEST LEVELS IN HEART, LIVER,  
 CC SKELETAL MUSCLE AND OVARY; LOWEST LEVELS IN BRAIN, LUNG AND  
 CC KIDNEY. ALSO FOUND IN B-CELLS, T-CELLS AND ENDOTHELIAL CELLS.  
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOKINE FAMILY OF RECEPTORS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; Y10659; CAA71669.1; -;  
 DR EMBL; Y09328; CAA70508.1; -;  
 DR EMBL; U62858; AAB37127.1; -;  
 DR HSSP; P31785; 11LN.  
 DR MIM; 300119; -;  
 DR PFAM; PF00041; fn3; 1.  
 KW Receptor; transmembrane; Glycoprotein; Immunoglobulin domain; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 427  
 FT DOMAIN 22 343  
 FT TRANSMEM 344 367  
 FT DOMAIN 368 427  
 FT DOMAIN 39 102  
 FT DISULFID 134 144  
 FT DISULFID 173 185  
 FT CARBOHYD 37 37  
 FT CARBOHYD 61 61  
 FT CARBOHYD 105 105  
 FT CARBOHYD 138 138  
 FT CARBOHYD 157 157  
 FT CARBOHYD 235 235  
 FT CARBOHYD 265 265  
 FT CARBOHYD 293 293  
 FT CARBOHYD 329 329  
 FT CARBOHYD 341 341  
 FT CONFLICT 130 130 T -> I (IN REF. 3).  
 FT CONFLICT 358 358 G -> D (IN REF. 3).  
 SQ SEQUENCE 427 AA; 48759 MW; 5983B3E8F554107B CRC64;

Query Match 5.8%; Score 90; DB 1; Length 427;  
 Best Local Similarity 30.9%; Pred. No. 6.28e-01;  
 Matches 21; Conservative 18; Mismatches 25; Indels 4; Gaps 4;  
 Db 122-EGDPESAVTELOCIQHNLSYMKCSWLPGRNTSPDTNVTLYVWHSLEKTHOCEN-IFREG 180  
 QY 23 QGGAEGV-QIQIYFNLETQVQTNWASKY-SR-TNLTFFHYRFGDEAYDQCTNVLQEQ 79  
 Db 181 QYFGCSFD 188  
 QY 80 HTSGCLLD 87  
 RESULT 15

US-09-376-430-2-05.rsp

Thu May 11 06:49:35 2000

```

ID   KS6B_XENLA  STANDARD;          PRT;   629 AA.
AC   P10666;
DT   01-JUL-1989 (Rel. 11, Created)
DI   01-JUL-1989 (Rel. 11, Last sequence update)
DE   15-DEC-1998 (Rel. 37, Last annotation update)
DE   RIBOSOMAL PROTEIN S6 KINASE II BETA (EC 2.7.1.1-) (S6KII-BETA)
DE   (P90-RSK).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC   Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae;
OC   Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE: 88217904.
RA   Jones S.W., Erikson E., Blenis J., Maller J.L., Erikson R.L.;
RI   "A Xenopus ribosomal protein S6 kinase has two apparent kinase
RT   domains that are each similar to distinct protein kinases.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:3377-3381(1988).
CC   !- FUNCTION: PHOSPHORYLATES A WIDE RANGE OF SUBSTRATES INCLUDING
CC   RIBOSOMAL PROTEIN S6. IMPLICATED IN THE ACTIVATION OF THE MITOGEN-
CC   ACTIVATED KINASE CASCADE.
CC   !- SIMILARITY: BELONGS TO THE PROTEIN KINASE SUPERFAMILY; SER/THR
CC   FAMILY. CONTAINS TWO KINASE DOMAINS.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; M20188; AAA49959.1; -.
DR   PIR; A30001; A30001.
DR   HSP; Q63450; IAO6.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 2.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 2.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 2.
DR   PFAM; PF00069; Pkinase; 2.
DR   PFAM; PF00433; Pkinase_C; 1.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Repeat; Multigene family.
FT   DOMAIN        62..321  PROTEIN KINASE 1.
FT   NP_BIND       68..76   PROTEIN KINASE 2.
FT   BINDING       94..94   ATP (BY SIMILARITY).
FT   ACT_SITE      187..187 ATP (BY SIMILARITY).
FT   NP_BIND       422..430 BY SIMILARITY.
FT   BINDING       445..445 ATP (BY SIMILARITY).
FT   ACT_SITE      533..533 BY SIMILARITY.
SQ   SEQUENCE 629 AA; 71286 MW; BC6144415FDECE0 CRC64;
Query Match          5.8%; Score 90; DB 1; Length 629;
Best Local Similarity 31.7%; Pred. NO. 6.28e-01;
Matches 19; Conservative 14; Mismatches 23; Indels 4; Gaps 3;
Db 475 DVEECNSIYLVTELMRGCELLDRILROK--FFSERACSVLFTYCK-TVEYLHSQGVVH 531
QY 66 EAYDOCTN-YLLQEGTSGCLLDRAEQRDILYFSIRNGTHFVFTASRWMMVYILKPSPKH 124

```

Search completed: Wed May 10 11:39:13 2000  
Job time : 108 secs.



\*\*\*\*\*  
MUSE  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:39:31 2000; MasPar time 239.74 Seconds  
Tabular output not generated. 58.708 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (23-225) from US09376430A.pep (5 of 25)  
Perfect Score: 1545  
Sequence: 1 OGGAAEGVQIIYFENLETV.....VTCWQGEIRDCAETPTTP 203

Scoring table: PAM 150  
Gap 11

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: sptrembl12  
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-mnc 8:sp-organelle  
9:sp-phase 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus

Statistics: Mean 42.839; Variance 70.120; scale 0.611

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Match	Length	ID	Description	Pred. No.
1	130	8.4	393	11	IL-13 RECEPTOR ALPHA 2	2.61e-07
2	109	7.1	435	2	PUTATIVE ASPARTATE AMI	1.26e-03
3	109	7.1	435	2	ASPARTATE AMINOTRANSFER	1.26e-03
4	100	6.5	890	11	INTERLEUKIN-5 RECEPTOR	3.70e-02
5	99	6.4	4436	1	4436AA LONG HYPOTHETIC	5.33e-02
6	99	6.4	6048	5	TWITCHIN	5.33e-02
7	99	6.4	6831	5	UNC-22 PROTEIN.	5.33e-02
8	99	6.4	7160	5	2K617.1B PROTEIN.	5.33e-02
9	97	6.3	311	2	CYTOKINE RECEPTOR COMP	1.10e-01
10	97	6.3	311	2	SIMILAR TO TRIMETHYLAM	1.10e-01
11	96	6.2	896	11	INTERLEUKIN-3 RECEPTOR	1.57e-01
12	95	6.1	315	11	MOR 5'BETA3.	2.23e-01
13	94	6.1	407	5	T15D6.9 PROTEIN.	3.18e-01
14	95	6.1	638	6	GROWTH HORMONE RECEPTO	2.23e-01
15	95	6.1	673	14	PUTATIVE RNA DEPENDENT	2.23e-01
16	94	6.1	919	5	PROBABLE EPH-LIKE KINA	3.18e-01
17	94	6.1	981	3	POSSIBLE DEHYDROGENEAS	3.18e-01
18	95	6.1	987	5	SIMILAR TO IMMUNOGLOBU	2.23e-01
19	94	6.1	1122	5	EPH RECEPTOR TYROSINE	3.18e-01
20	94	6.1	26926	4	TITIN, HEART ISOFORM N	3.18e-01

21 93 6.0 173 5 Q18307 COSMID C29F5. 4.51e-01  
22 92 6.0 202 14 P90279 NEF PROTEIN. 6.37e-01  
23 92 6.0 684 13 Q05444 NOVEL ANTIGEN RECEPTOR 6.37e-01  
24 92 6.0 1385 5 Q4924 ROUNDABOUT 1. 8.99e-01  
25 91 5.9 202 14 P88435 NEF PROTEIN. 1.26e-00  
26 90 5.8 204 14 Q74519 NEF. 1.26e-00  
27 90 5.8 204 14 Q74920 NEF. 1.26e-00  
28 90 5.8 204 14 Q74921 NEF. 1.26e-00  
29 89 5.8 255 2 Q9X207 CONSERVED HYPOTHETICAL 1.77e-00  
30 90 5.8 255 10 Q65756 VEGETATIVE LECTIN. KIN 1.77e-00  
31 89 5.8 339 2 Q92199 PUTATIVE HISTIDINE KIN 1.26e-00  
32 90 5.8 376 1 Q9Y8P6 376AA LONG HYPOTHETICA 1.26e-00  
33 90 5.8 427 4 Q95646 INTERLEUKIN-13 RECEPTO 1.26e-00  
34 89 5.8 634 6 Q46600 SOMATOTROPIN RECEPTOR 1.77e-00  
35 90 5.8 935 2 Q31000 INTIMIN. 1.26e-00  
36 89 5.8 1286 10 Q92R72 P-GLYCOPROTEIN. 1.77e-00  
37 89 5.8 1466 3 Q42930 VACUOLAR PROTEIN SORTI 1.77e-00  
38 88 5.7 137 14 Q07344 STRUCTURAL PROTEIN P11 2.48e-00  
39 88 5.7 269 6 Q79195 GROWTH HORMONE BINDING 2.48e-00  
40 88 5.7 366 3 Q42858 HYPOTHETICAL 41.7 KD T 2.48e-00  
41 88 5.7 459 2 Q86960 ALPHA-GLUCOSIDASE. 2.48e-00  
42 88 5.7 934 2 Q85627 L0025. 2.48e-00  
43 87 5.6 351 5 Q96548 DIAMINOPIMELATE EPIMER 3.46e-00  
44 87 5.6 461 4 Q75072 FUKUTIN. 3.46e-00  
45 87 5.6 563 14 Q9WT04 GANCICLOVIR KINASE, CO 3.46e-00

ALIGNMENTS

RESULT 1  
ID O88786 PRELIMINARY; PRT; 383 AA.  
AC O88786;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DE 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE IL-13 RECEPTOR ALPHA 2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
RN Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RP SEQUENCE FROM N.A.  
RC STRAIN=C3H/HEJ; TISSUE=THYMUS;  
RX MEDLINE: 98391042.  
RA DONALDSON D.D., WHITTERS M.J., FITZ L., NEBEN T.Y., FINNERTY H.,  
RA HENDERSON S.L., O'HARA R.M. JR., BEIER D.R., TURNER K.J., WOOD C.R.,  
RA COLLINS M.,  
RT "The murine IL-13 receptor alpha 2: molecular cloning,  
characterization, and comparison with murine IL-13 receptor alpha 1.";  
RL J. Immunol. 161:2317-2324(1998).  
DR EMBL: U65747; AAC33240.1; -;  
SQ SEQUENCE 383 AA: 44483 MW: 5EAEF3E3 CRC32;

Query Match 8.4%; Score 130; DB 11; Length 383;  
Best Local Similarity 30.2%; Pred. No. 2.61e-07;  
Matches 26; Conservative 22; Mismatches 32; Indels 6; Gaps 5;  
Db 127 EGSLETKIQDMKCIYNNWQYLVCSWPKGTIVSDTNTMFFWYEGLDHAL-QCADIQLQHD 185  
QY 23 OGGAAEGVQ-IQIIYFENLETVQVTWNASK--YSRTNLTHYRENG-DEAYDOCTNYLLQE 78  
Db 186 EKNVCKLSNLDSSDYKDFFCVNGS 211  
QY 79 GHTSGCLLDAAEQRDIL-YFSIRNGT 103

RESULT 2  
ID O06885 PRELIMINARY; PRT; 435 AA.  
AC O06885;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DE 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE PUTATIVE ASPARTATE AMINOTRANSFERASE TPAAT.  
GN TPAAT.

```

OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PALLIDUM.
RX MEDLINE; 97294457.
RA SHEVCHENKO D.V., AKINS D.R., ROBINSON E., LI M., POPOVA T.G.,
RA COX D.L., RADOLF J.;
RT "Molecular characterization and cellular localization of TpLRR, a
RT processed leucine-rich repeat protein of Treponema pallidum, the
RT syphilis spirochete.";
RL J. Bacteriol. 179:3188-3195(1997).
DR EMBL; U73748; AAC45303.1; -.
KW Transferase; Aminotransferase.
SQ SEQUENCE 435 AA; 47285 MW; 7AF01306 CRC32;

Query Match          7.1%; Score 109; DB 2; Length 435;
Best Local Similarity 35.1%; Pred. No. 1.26e+03;
Matches 20; Conservative 11; Mismatches 21; Indels 5; Gaps 5;

Db    206 AVLVICDD-AYSGEFYEASLMRGSFFARFAQAHNKICALKIDGLTKEE-YA-WGLRV 259
      |||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||..|||
Qy    133 AVTVTCDSLYGDLLEYEQY-RSPFDWFQSKQNTCNVTIEGLDAEKCYFW-VRV 187

RESULT 3
ID O83252 PRELIMINARY; PRT; 435 AA.
AC O83252;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE ASPARTATE AMINOTRANSFERASE (TPAAT).
GN TP0223.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98332770.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.,
RA DODSON R., GWINN M., HICKY E.K., CLAYTON R., KETCHUM K.A.,
RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.,
RA McDONALD L., ARTTACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
RA HATCH B., HORST K., ROBERTS K., WATHEY L., WEIDMAN J., SMITH H.O.,
RA VENTER J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
[2]
SE SEQUENCE FROM N.A.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.,
RA DODSON R., GWINN M., HICKY E.K., CLAYTON R., KETCHUM K.A.,
RA SODERGREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.,
RA McDONALD L., ARTTACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
RA HATCH B., HORST K., ROBERTS K., WATHEY L., WEIDMAN J., SMITH H.O.,
RA VENTER J.C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE001204; AAC65212.1; -.
DR TIGR; TP0223; -.
SQ SEQUENCE 435 AA; 47304 MW; 0B31A6F1 CBC32.
KW Transferase; Aminotransferase.

```

```

Query Match          7.1%; Score 109; DB 2: Length 435;
Best Local Similarity 35.1%; Pred. No. 1,26e-03;
Matches 20; Conservative 11; Mismatches 21; Indels 5; Gaps 5;
Db 206 AVLVICDD-AYSGFEYEASLMRGSTAFRAQAHHKNCALKDGLTRKEE-YA-WGLVR 259
QY 133 AVTVCTSDLSYGDLLYEQY-RSPDETQWSQKQENTCNVTIEGLDAKCVSPW-VRV 187

RESULT 4

```

ID	Q9Z1AO		PRELIMINARY;	PRT;	890 AA.					
AC	Q9Z1AO;									
DT	01-MAY-1999	(TREMBRel. 10,	Created)							
DT	01-MAY-1999	(TREMBRel. 10,	Last sequence update)							
DT	01-NOV-1999	(TREMBRel. 12,	Last annotation update)							
DE	INTERLEUKIN-5 RECEPTOR BETA CHAIN.									
GW	IL5.									
OS	Cavia porcellus (Guinea pig).									
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;									
CC	Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RA	LOGSDON N.J., GRAHAM A., SCOTT C.W.;									
RT	"Guinea pig IL5 receptor beta chain."									
RL	Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.									
DR	EMBL; U94686; AAC77520.1; -.									
DR	HSSP; P40189; IBQU.									
KW	Receptor.									
SQ	SEQUENCE	890 AA;	96579 MW;	EABC43CC CRC32;						
						Query Match	6.5%;	Score 100;	DB 11;	Length 890;
						Best Local Similarity	34.3%;	Pred. No. 3,70e-02;		
						Matches	23;	Conservative	13;	Mismatches 23; Indels 8; Gaps 6;
Dd	375	FELIQRTAGGR-WENSKETLKNAHNNPLPLEPATYTLARVKVSPGAGYN-GTW-SE	431							
QY	148	YEVOYRSPFDTTEWQ-SKQE---NTCNVTIEGLDAERKCYSFWVRVKAM-EDVYGFTYPSD	202							
Dd	432	WSBEQRW	438							
QY	203	WSEVTCW	209							

RESULT	5	
ID	058659	PRELIMINARY; PRT; 4436 AA.
AC	058659;	
DT	01-AUG-1998 (TrEMBLrel. 07, Created)	
DT	01-AUG-1998 (TrEMBLrel. 07, Last sequence update)	
DT	01-MAY-1999 (TrEMBLrel. 10, Last annotation update)	
DE	4436AA. LONG HYPOTHETICAL PROTEIN.	
GN	PH0954.	
OS	Pyrococcus horikoshii.	
OC	Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.	
RN	[1]	
RC	SEQUENCE FROM N.A.	
RC	STRAIN-OT3.	
RC	MEDLINE; 98344137.	
RA	KAWABAYASI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,	
RA	YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOIYAMA A., NAGAI Y.,	
RA	SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,	
RA	FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,	
RA	AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,	
RA	KIKUCHI H.;	
RT	"Complete sequence and gene organization of the genome of a hyper-	
RT	thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";	
RL	DNA Res. 5:55-76(1998).	
RL	EMBL; AP000004; BAA30051.1; -.	
DR	PFAM; PF00041; fn3; 1.	
DR	PFAM; PF00801; PRD; 2.	
SQ	SEQUENCE 4436 AA; 497015 MW; 2D9608CA CRC32;	

```
Query Match          .6.4%;      Score 98;   DB 1; Length 4436;
Best Local Similarity 26.8%;      Pred. N. 5.33e-02;
Matches 19; Conservative 18; Mismatches 29; Indels 5; Gaps 5;

Ddb    4115  NMTYNETIREIKVRAT-D-E-GIANVTATINGSLSEKVNWTWIGRVELDDGKYELNV 4171
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     140  DLSYGQLLYEVOYRSPDETQSQKENTCNVTIEGLDAEKCYSEWV-RVKAMEDVIYGPDT 198
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Ddb    4172  FASDKGNNGVC 4182
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     199  YPSD-WSEVTC 208
       : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```



RESULT 6 PRELIMINARY; PRT; 6048 AA.  
 ID Q23020 Q27232;  
 AC Q23020; Q27232; (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)  
 DE TWITCHIN.  
 GN UNC-22 OR ZK617.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 90044042.  
 RA BENJIAN G.M., KIFF J.E., NECKELMANN N., MOERMAN D.G., WATERSON R.H.;  
 RT "Sequence of an unusually large protein implicated in regulation of  
 RT myosin activity in C. elegans.";  
 RL Nature 342:45-50(1989).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 93387664.  
 RA BENJIAN G.M., L'HERNAULT S.W., MORRIS M.E.;  
 RT "Additional sequence complexity in the muscle gene, unc-22, and its  
 RT encoded protein, twitchin, of Caenorhabditis elegans.";  
 RL Genetics 134:1097-1104(1993).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RA HARRIS B.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X15423; CAA3463.1; -;  
 DR EMBL; Z73899; CAA98081.1; ALT\_INT.  
 DR EMBL; Z73897; CAA98081.1; JOINED.  
 DR HSP; Q63450; IAO6.  
 DR PFAM; PF00041; fn3; 31.  
 DR PFAM; PF00047; ig; 13.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PRINTS; PR00014; FNTYPEIII.  
 DR MYOSIN; Kinase.  
 KW MYOSIN; Kinase.  
 SQ SEQUENCE 6048 AA; 668449 MW; 1977C602 CRC32;  
 Query Match 6.4%; Score 99; DB 5; Length 6048;  
 Best Local Similarity 33.9%; Pred. No. 5.33e-02;  
 Matches 19; Conservative 8; Mismatches 27; Indels 2; Gaps 2;  
 Db 916 YIVEVRDPDTKEWKEVKRVPDTNASISGLKEGKEYQFRVAVNKAGP-GQPSEPE 970  
 QY 148 YEVQIRSPFTEW-QSKQENTCNVTIEGLDAEKCYFWRVKAMEDVYGPDTYPSD 202  
 RESULT 7 PRELIMINARY; PRT; 6831 AA.  
 ID Q23550  
 AC Q23550;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)  
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)  
 DE UNC-22 PROTEIN.  
 GN UNC-22  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA WHITE S.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z73897; CAA98064.1; -;  
 DR EMBL; Z73899; CAA98064.1; JOINED.  
 DR HSP; F02751; IFNA.  
 DR PFAM; PF00041; fn3; 31.  
 DR PFAM; PF00047; ig; 17.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PRINTS; PR00014; FNTYPEIII.  
 DE UNC-22  
 GN UNC-22  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA WHITE S.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z73897; CAA98064.1; -;  
 DR EMBL; Z73899; CAA98064.1; JOINED.  
 DR HSP; F02751; IFNA.  
 DR PFAM; PF00041; fn3; 31.  
 DR PFAM; PF00047; ig; 17.  
 DR PFAM; PF00069; pkinase; 1.

DR PRINTS; PR00014; FNTYPEIII.  
 SQ SEQUENCE 6831 AA; 752579 MW; 0A66C338 CRC32;  
 Query Match 6.4%; Score 99; DB 5; Length 6831;  
 Best Local Similarity 33.9%; Pred. No. 5.33e-02;  
 Matches 19; Conservative 8; Mismatches 27; Indels 2; Gaps 2;  
 Db 1699 YIVEVRDPDTKEWKEVKRVPDTNASISGLKEGKEYQFRVAVNKAGP-GQPSEPE 1753  
 QY 148 YEVQIRSPFTEW-QSKQENTCNVTIEGLDAEKCYFWRVKAMEDVYGPDTYPSD 202  
 RESULT 8 PRELIMINARY; PRT; 7160 AA.  
 ID Q23551  
 AC Q23551; 1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 08, Last sequence update)  
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)  
 DE ZK617.1B PROTEIN.  
 GN ZK617.1B.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA HARRIS B.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPESE T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LAITREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKIN L., SAUNDERS D., SAUNDERS R.,  
 RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATKINSON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RA WHITE S.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z73899; CAA98082.1; -;  
 DR EMBL; Z73897; CAA98082.1; JOINED.  
 DR EMBL; Z73897; CAA98065.1; -;  
 DR EMBL; Z73899; CAA98065.1; JOINED.  
 DR HSP; P02751; IFNA.  
 DR PFAM; PF00041; fn3; 31.  
 DR PFAM; PF00047; ig; 17.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PRINTS; PR00014; FNTYPEIII.  
 SQ SEQUENCE 7160 AA; 789211 MW; EDD567FE CRC32;  
 Query Match 6.4%; Score 99; DB 5; Length 7160;  
 Best Local Similarity 33.9%; Pred. No. 5.33e-02;  
 Matches 19; Conservative 8; Mismatches 27; Indels 2; Gaps 2;  
 Db 2028 YIVEVRDPDTKEWKEVKRVPDTNASISGLKEGKEYQFRVAVNKAGP-GQPSEPE 2082  
 QY 148 YEVQIRSPFTEW-QSKQENTCNVTIEGLDAEKCYFWRVKAMEDVYGPDTYPSD 202  
 RESULT 9 PRELIMINARY; PRT; 256 AA.  
 ID Q63968  
 AC Q63968;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1999 (TREMblrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)

```

DE  CYTOKINE RECEPTOR COMPLEX COMMON BETA CHAIN H BETA C (FRAGMENT).
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC  Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE; 94235843.
RA  D'ANDREA R., RAYNER J., MORETTI P., LOPEZ A., GOODALL G.J.,
RA  GONDA T.J., VADAS M.;
RT  "A mutation of the common receptor subunit for interleukin-3 (IL-3),
RT  granulocyte-macrophage colony-stimulating factor, and IL-5 that leads
RT  to ligand independence and tumorigenicity.";
RL  Blood 83:2802-2808(1994).
DR  EMBL; S70302; AAB31055.1; -.
DR  HSSP; P40189; IBOU.
DR  PFAM; PF00041; fn3; 1.
FT  NON_TER 1
SQ  SEQUENCE 256 AA; 29206 MW; 5242B76B CRC32;

Query Match 6.3%; Score 97; DB 11; Length 256;
Best Local Similarity 22.6%; Pred. No. 1.10e-01;
Matches 14; Conservative 16; Mismatches 29; Indels 3; Gaps 3;

Db 80 RTGYNIGWSESEAH-SMALPALEPSTRYWARVRVTSRTGYN-GIW-SENSEARSWDTE 136
QY 153 RSPDTEWQSKQNTCNVTIEGLDAEKCYSFWRVKAMEDVGPDPYSDWSEVTCWORG 212
Db 137 SV 138
QY 213 EI 214

RESULT 10
ID Q48962 PRELIMINARY; PRT; 311 AA.
AC Q48962;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE SIMILAR TO TRIMETHYLAMINE DH (FRAGMENT).
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Bacillus/clostridium group; Mollicutes;
OC capricolum group.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 27343(KID);
RX MEDLINE; 96059841.
RA BORK P., OUFOUNIS C., CASARI G., SCHNEIDER R., SANDER C., DOLAN M.,
RA GILBERT W., GILLEVET P.M.;
RT "Exploring the Mycoplasma capricolum genome: a minimal cell reveals
RT its physiology.";
RL Mol. Microbiol. 16:955-967(1995).
DR EMBL; Z33015; CAA83700.1; -.
DR PFAM; PF00724; oxidored_FMN; 1.
FT NON_TER 311
SQ SEQUENCE 311 AA; 35768 MW; 6CC72E66 CRC32;

Query Match 6.3%; Score 97; DB 2; Length 311;
Best Local Similarity 27.9%; Pred. No. 1.10e-01;
Matches 12; Conservative 11; Mismatches 19; Indels 1; Gaps 1;

Db 202 FCLEVYKAIREVDKYPKNFIFGFRATPEETGYDILGYTIED 244
QY 37 FNLETVQVWNA-SKYSRTNLTHYRNGEAYDQCTNYLLOE 78

RESULT 11
ID Q64146 PRELIMINARY; PRT; 896 AA.
AC Q64146;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE INTERLEUKIN-3 RECEPTOR BETA-SUBUNIT (FRAGMENT).
GN RIL-3R<BETA>.

```

```

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95370942.
RA APPEL K., BUTTINI M., SAUTER A., GEBICKE-HARTER P.J.;
RT "Cloning of rat interleukin-3 receptor beta-subunit from cultured
RT microglia and its mRNA expression in vivo.";
RL J. Neurosci. 15:5800-5809(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-WISTAR; TISSUE-BRAIN;
RA GEBICKE-HARTER P.J.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; S79263; AAB35068.1; -.
DR EMBL; AT000555; CAA04186.1; -.
DR PFAM; PF00041; fn3; 2.
KW Signal.
FT NON_TER 1
SQ SEQUENCE 896 AA; 99504 MW; 8E7ED2CD CRC32;

Query Match 6.2%; Score 96; DB 11; Length 896;
Best Local Similarity 30.3%; Pred. No. 1.57e-01;
Matches 20; Conservative 17; Mismatches 21; Indels 8; Gaps 8;

Db 375 FOVQYKKLDR-WEDSKTENLNAHSMDLPLQLEPGTGYCARVRVKTPIE-Y-KGLW-SEW 430
QY 148 YEYQYRSPDTEWQ-SKQNTCNV-TIE-G-LDAEKCYSFWRVKAMEDVGPDPYSDW 203
Db 431 SNECTW 436
QY 204 SEVTCW 209

RESULT 12
ID Q9WVN6 PRELIMINARY; PRT; 315 AA.
AC Q9WVN6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE MOR 5'BETA3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-129;
RX MEDLINE; 99055560.
RA BENDER M.A., REIK A., CLOSE J., TELLING A., EPNER E., FIERING S.,
RA HARDISON R., GROUNDINE M.;
RT "Description and targeted deletion of 5' hypersensitive site 5 and 6
RT of the mouse beta-globin locus control region.";
RL Blood 92:4394-4403(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-129;
RX MEDLINE; 99238494.
RA BULGER M., VON DOORNICK J.H., SAITOH N., TELLING A., FARRELL C.,
RA BENDER M.A., FELSENFELD G., AXEL R., GROUNDINE M.;
RT "Conservation of sequence and structure flanking the mouse and human
RT beta-globin loci: the beta-globin genes are embedded within an array
RT of odorant receptor genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:5129-5134(1999).
DR EMBL; AF071080; AAD28304.1; -.
SQ SEQUENCE 315 AA; 35559 MW; DEE53BF7 CRC32;

Query Match 6.1%; Score 95; DB 11; Length 315;
Best Local Similarity 28.3%; Pred. No. 2.23e-01;
Matches 13; Conservative 13; Mismatches 17; Indels 3; Gaps 3;

Db 159 PIIRLHWFPYC-RSHVLSHA-FCLHQDVIKLACADITF-NRLYPV 201

```

US-09-376-430-2-05.rspt

Thu May 11 06:49:36 2000

Matches 16; Conservative 12; Mismatches 22; Indels 2; Gaps 2;

Db 190 LEYELQYKEVNETKMKMDPILSTSPVYSLKVDKEYEVVRVSRKRRNSGNYG 241  
 QY 146 LLYEVQIRSPFDTEWOSKQEN-TCNVITIEGLDAEKCYSFVVRVKAMEDV-YG 195

RESULT 15  
 ID Q9YX75 PRELIMINARY; PRT; 673 AA.  
 AC Q9YX75;  
 DT 01-MAY-1999 (Tremblrel. 10, Created)  
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)  
 DE PUTATIVE RNA DEPENDENT RNA POLYMERASE.  
 OS Fusarium poae virus 1 (FUP0-1).  
 OC Viruses; dsRNA viruses; Partitiviridae; Partitivirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-All;  
 RA COMPEL P.; PAPP I.; BIBO M.; FEKETE C.; HORNOK L.;  
 RT "Genetic interrelationships and genome organization of double-stranded  
 RNA elements of Fusarium poae.";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF047013; AAC98734.1;  
 SQ SEQUENCE 673 AA; 78315 MW; C04FA621 CRC32;

Query Match 6.1%; Score 95; DB 14; Length 673;  
 Best Local Similarity 29.4%; Pred. No. 2.23e-01;  
 Matches 25; Conservative 21; Mismatches 30; Indels 9; Gaps 8;

Db 573 YMSARVGMANASCGQDFTFDFCDVYHEFNDRADLDE-SAYLHIOHLPGLKIDES 631  
 QY 36 YFNLETVQVTNNASKYSRINTLF-H-YR-FNGDEA-YDOCTNYLLOEGHTSGCL-LDAE 89  
 Db 632 VRQ-IVDFQVFPSPQQTVYHTVSRWK 655  
 QY 90 QRDDILYFSIRNGTHPVF-TASRW 113

Search completed: Wed May 10 11:43:41 2000  
 Job time : 250 secs.

QY 105 PVTASRWVYILKPSPKHVRFSWHDQAVTVTCSDLSYGDLLYEV 150

RESULT 13  
 ID Q9XU71 PRELIMINARY; PRT; 407 AA.

AC Q9XU71;  
 DT 01-NOV-1999 (Tremblrel. 12, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
 DE T15D6.9 PROTEIN.  
 GN T15D6.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE; 94150718.  
 RA DOBSON R.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.

RX MEDLINE; 94150718.  
 RA WILSON R.; AINSOUGH R.; ANDERSON K.; BAYNES C.; BERKS M.;  
 RA BONFIELD J.; BURTON J.; CONNELL M.; COPSEY T.; COOPER J.; COULSON A.;  
 RA CRAXTON M.; DEAR S.; DU Z.; DURBIN R.; FAVELLO A.; FULTON L.;  
 RA GARDNER A.; GREEN P.; HAWKINS T.; HILLIER L.; JIER M.; JOHNSTON L.;  
 RA JONES M.; KERSHAW J.; KIRSTEN J.; LAISTER N.; LATREILLE P.;  
 RA LIGHTNING J.; LLOYD C.; MCMURRAY A.; MORTIMORE B.; O'CALLAGHAN M.;  
 RA PARSONS J.; PERCY C.; RIFKEN L.; ROOPRA A.; SAUNDERS D.; SHOWNKEEN R.;  
 RA SMAILDON N.; SMITH A.; SONNHAMMER E.; STADEN R.; SULTON J.;  
 RA THERRY-MIEG J.; THOMAS K.; VAUDIN M.; VAUGHAN K.; WATERSTON R.;  
 RA WATSON A.; WEINSTOCK L.; WILKINSON-SPOAT J.; WOLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; 283125; CAB05619.1;  
 SQ SEQUENCE 407 AA; 47157 MW; 1487A379 CRC32;

Query Match 6.1%; Score 94; DB 5; Length 407;  
 Best Local Similarity 30.4%; Pred. No. 3.18e-01;  
 Matches 17; Conservative 15; Mismatches 20; Indels 4; Gaps 4;

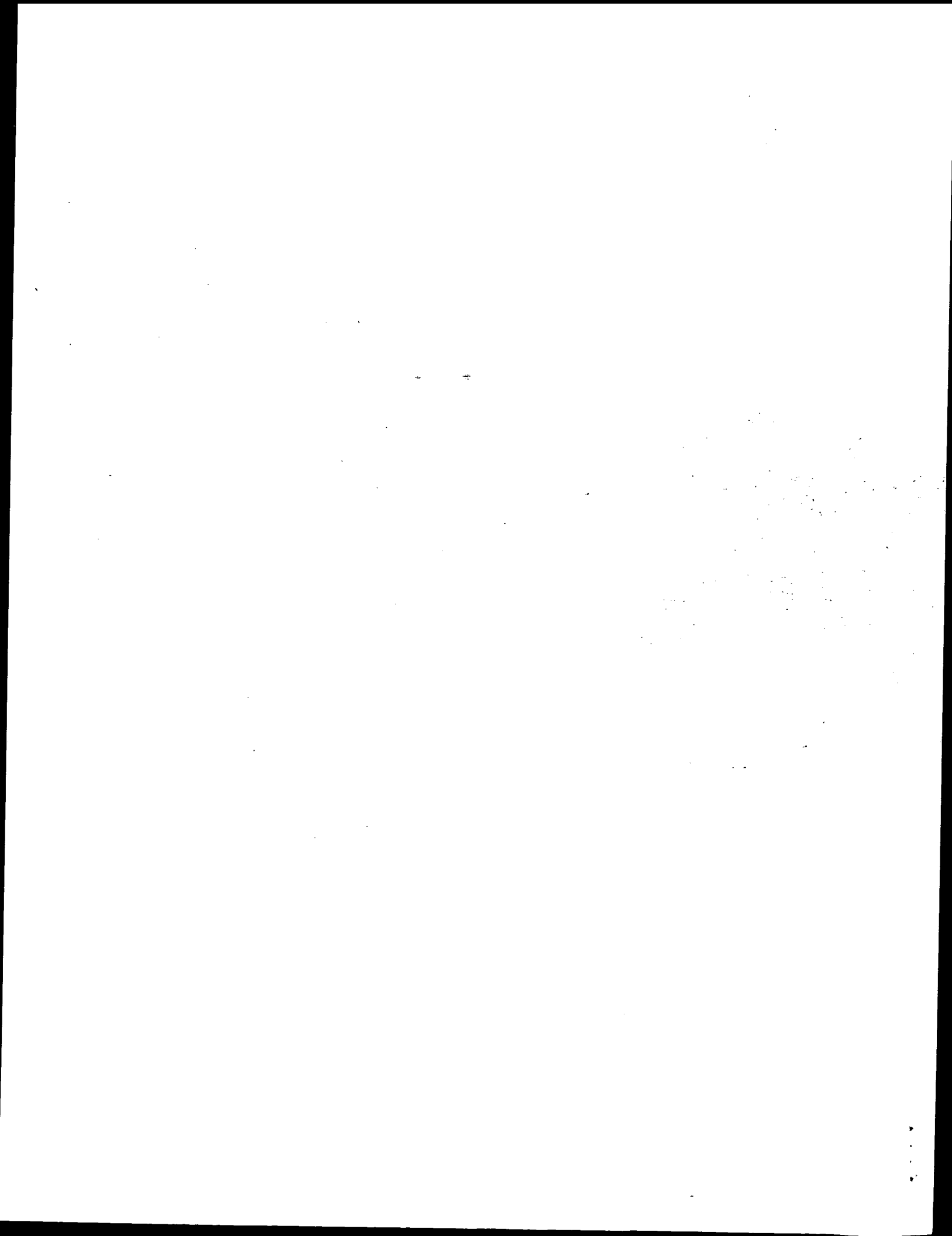
Db 117 EYFPH-HKYLHGLNLSAAVIDKLPRNDEFRAFTSTYKPVSN-WMTYNFKP 170  
 QY 66 EAYDQCTNYLLOEGHTSGCLLDAEQRDDILY-F-SIRNGTHPVFTASRWVYILKP 119

RESULT 14  
 ID Q9XSZ1 PRELIMINARY; PRT; 638 AA.

AC Q9XSZ1;  
 DT 01-NOV-1999 (Tremblrel. 12, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
 DE GROWTH HORMONE RECEPTOR.  
 OS Papio anubis (olive baboon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;  
 OC Papio.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA ZOGOPoulos G.; NATHANIELSZ P.; HENDY G.N.; GOODYER C.G.;  
 RT "The baboon: a model for the study of primate growth hormone receptor  
 gene expression during development.";  
 RL J. Mol. Endocrinol. 23:0-0(1999).  
 DR EMBL; AF150751; AAD39536.1;  
 KW Receptor.  
 SQ SEQUENCE 638 AA; 71407 MW; 2EC386D7 CRC32;

Query Match 6.1%; Score 95; DB 6; Length 638;  
 Best Local Similarity 30.8%; Pred. No. 2.23e-01;



\*\*\*\*\*  
 M A S E R H  
 \*\*\*\*\*  
 (TH)

Release 3.1A John F. Collins, BioComputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 11:52:56 2000; MasPar time 2.87 Seconds

Tabular output not generated. 57.811 Million cell updates/sec

Title: >US-09-376-430-2

Description: (198-204) from US09376430A.pap (6 of 25)

Sequence: 1 TYPEDWS 7

Scoring table: PAM 150

Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: a-geneseq35

i:geneseqp

Statistics: Mean 16.717; Variance 56.206; scale 0.297

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	ID	Description	Pred. No.
1	52	91.2	1 W05411	Human H74 protein.	1.09e+02
2	49	85.0	2 W29878	Lysophosphatidic acid	2.14e+02
3	48	84.2	3 W28506	Ad4/Ad3LP sequence.	2.67e+02
4	47	82.5	4 W39829	HIV Vpr protein sequen	3.33e+02
5	47	82.5	5 W98811	HIV E21.24P protein se	3.33e+02
6	47	82.5	6 W98812	HIV-1 Vpr protein.	3.33e+02
7	47	82.5	7 W39504	Streptococcus pneumoni	3.33e+02
8	47	82.5	8 W05410	Mouse H74 protein.	3.33e+02
9	46	80.7	9 W13095	Trypanosoma cruzi anti	4.15e+02
10	46	80.7	10 W42395	Pyrococcus furiosus VC	4.15e+02
11	46	80.7	11 W1970	C. elegans CEF37C12.2	4.15e+02
12	46	80.7	12 R13615	Protein found during G	4.15e+02
13	46	80.7	13 R93090	Human interleukin-11 r	4.15e+02
14	46	80.7	14 R92814	Human interleukin-11 r	4.15e+02
15	46	80.7	15 W1562	Human acetylcholine re	4.15e+02
16	46	80.7	16 R73751	Mouse S22L protein.	4.15e+02
17	46	80.7	17 W98830	Thermotoga thermophil	4.15e+02
18	45	78.9	18 W98830	HIV Vpr protein sequen	5.16e+02
19	45	78.9	19 W98833	HIV A30L protein sequen	5.16e+02
20	45	78.9	20 W98834	HIV A59P protein sequen	5.16e+02
21	45	78.9	21 W98838	HIV G76S protein sequen	5.16e+02
22	45	78.9	22 W98837	HIV G75A protein sequen	5.16e+02
23	45	78.9	23 W98835	HIV L68S protein sequen	5.16e+02

24 45 78.9 96 1 W99836 HIV H71C protein seque 5.16e+02  
 25 45 78.9 96 1 W99832 HIV A30S protein seque 5.16e+02  
 26 45 78.9 119 1 R24555 Human x mouse modified 5.16e+02  
 27 45 78.9 119 1 R24560 Human x mouse modified 5.16e+02  
 28 45 78.9 119 1 P0540 Immunoglobulin H chain 5.16e+02  
 29 45 78.9 119 1 R24556 Human x mouse modified 5.16e+02  
 30 45 78.9 119 1 R24562 Human x mouse modified 5.16e+02  
 31 45 78.9 119 1 R24557 Human x mouse modified 5.16e+02  
 32 45 78.9 119 1 R24558 Human x mouse modified 5.16e+02  
 33 45 78.9 119 1 R24561 Human x mouse modified 5.16e+02  
 34 45 78.9 138 1 P0542 Amino acid sequence of 5.16e+02  
 35 45 78.9 199 1 W15774 Protein cognate of pro 5.16e+02  
 36 45 78.9 273 1 R44157 Moraxella bovis mboC g 5.16e+02  
 37 45 78.9 278 1 R62519 Isoform lg11 of the ga 5.16e+02  
 38 45 78.9 279 1 W00088 RAS-related protein en 5.16e+02  
 39 45 78.9 279 1 R14847 Protein associated wit 5.16e+02  
 40 45 78.9 432 1 R92813 Murine interleukin-11 5.16e+02  
 41 45 78.9 441 1 R99091 Murine Etl-2 gene prod 5.16e+02  
 42 45 78.9 702 1 W20286 H. pylori outer membra 5.16e+02  
 43 45 78.9 797 1 W20993 H. pylori outer membra 5.16e+02  
 44 45 78.9 872 1 W75912 Helicobacter valyl-trn 5.16e+02  
 45 45 78.9 874 1 W98698 H. pylori GNP0 686 pro 5.16e+02

## ALIGNMENTS

RESULT 1  
 ID W05411 standard; Protein; 377 AA.  
 AC W05411; 1998 (first entry)  
 DT 23-FEB-1998  
 DE Human H74 protein.  
 KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;  
 KW cellular signalling element; cellular structural element; malignancy;  
 KW protein identification; functional domain; protein screening;  
 KW cellular signal transduction process.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Misc\_difference 1 /note= "encoded by GA"  
 FT  
 PN W09631625-A1.  
 PD 10-OCT-1996.  
 PE 04-APR-1996; U04454.  
 PR 03-APR-1996; US-630915.  
 PR 07-APR-1995; US-417872.  
 PA (CYTO-) CYTOGEN CORP.  
 PA (CYNC-) UNIV NORTH CAROLINA.  
 PI Fowlkes DM, Hoffman N, Kay BK, McConnell SJ, Sparks AB;  
 DR WPI; 96-465045/46.  
 DR N-PSDB; T39810.  
 PT Identifying polypeptide(s) having specific functional domain (esp.  
 PT SH3 domain) - comprises detecting selective binding to recognition  
 PT unit, regardless of sequence homology  
 PS Claim 102; Fig 49; 174pp; English.  
 CC W05405-W05411 represent human and mouse Src-homology region 3 (SH3)  
 CC domain containing proteins that can be used in the method of the  
 CC invention. SH3 domain containing proteins play a role in signalling and  
 CC structural elements of cells. The method of the invention is for  
 CC identifying polypeptides containing functional domains of interest  
 CC (especially SH3 domains). The method comprises contacting a multivalent  
 CC recognition unit (RU) complex with a number of peptides and identifying  
 CC polypeptides having a selective binding affinity for the RU complex. The  
 CC method is based on functional similarities and does not rely on sequence  
 CC similarities. Prior methods only gave limited success for identifying  
 CC proteins which contain an SH3 domain due to the minimal sequence homology  
 CC among known SH3 proteins. It has been found that small peptide RUS in  
 CC multivalent form have reduced specificity for a given functional domain  
 CC compared to monomer RUS. Multivalent RU complexes are particularly suited  
 CC to screening for polypeptides in sequence to, the original target  
 CC similar to, but not identical in function to, the original target  
 CC functional domain. The new method enables proteins having a common  
 CC function to be identified. Identification of novel SH3 proteins will be  
 CC useful for a better understanding of cell growth, malignancy, signal  
 CC transduction processes, etc. New candidate drugs can be identified, and

CC their specificities (e.g. pharmacological activities) can be assessed  
 CC using the method of the invention.  
 SQ Sequence 377 AA;

Query Match 91.2%; Score 52; DB 1; Length 377;  
 Best Local Similarity 71.4%; Pred. No. 1.09e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 284 SYPTDWS 290  
 QY 198 YPDSWS 204  
 :||:|

RESULT 2  
 ID W29878 standard; Protein; 367 AA.  
 AC W29878;  
 DT 18-MAR-1998 (first entry)  
 DE Lysophosphatidic acid acyltransferase enzyme LPAAT I splice variant 2.  
 KW Lysophosphatidic acid acyl transferase; LPAAT; mammalian;  
 OS Homo sapiens; signal transduction.  
 FH Key Location/Qualifiers  
 FT Misc\_difference 37 /note= "X is encoded by a STOP codon"  
 FT  
 PN RD-400054-A.  
 PD 10-AUG-1997.  
 PF 20-JUL-1997; 400054.  
 PR 20-JUL-1997; RD-400054.  
 PA (ELMO/) ELMORE M A.  
 PA (FINN/) FINNEN M J.  
 PA (HILL/) HILL M E.  
 PA (KEL/) KELLY K.  
 PA (MAKD/) MAKDA A A.  
 PA (STAM/) STAMPS A.  
 PA (YAMA/) YAMANOUCHI RES INST.  
 PI Elmore MA, Finnen MJ, Hill ME, Kelly K, Makda AA,  
 PI Stamps A;  
 DR WPI; 97-433268/40.  
 DR N-PSDB; T85931.  
 PT Mammalian lyso-phosphatidic acid acyl-transferase enzymes - and  
 PT related DNA, useful for isolating inhibitors and studying the role  
 PT of the enzymes in signal transduction  
 PS Disclosure; Page -; 5pp; English.  
 CC The present sequence represents a novel mammalian lysophosphatidic acid  
 CC acyltransferase (LPAAT) enzyme, designated LPAAT I. Enzymes LPAAT I,  
 CC LPAAT II and LPAAT III are human homologues of non mammalian forms of  
 CC LPAAT. The sequences can be used to screen for LPAAT inhibitors and to  
 CC study the role of LPAAT enzymes in signal transduction and disease.  
 CC note: the present sequence does not appear in the specification; it was  
 CC created using information provided.  
 SQ Sequence 367 AA;

Query Match 86.0%; Score 49; DB 1; Length 367;  
 Best Local Similarity 66.7%; Pred. No. 2.14e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 78 YPSEWA 83  
 QY 199 YPDSWS 204  
 :||:|

RESULT 3  
 ID W28506 standard; Protein; 376 AA.  
 AC W28506;  
 DT 07-DEC-1997 (first entry)  
 DE AD4/AD3LP sequence.  
 KW AD3; AD4/AD3IP; Alzheimer's disease; chromosome; missegregation;  
 KW Presenilin; inhibitor; trisomy 21; AD.  
 OS Homo sapiens.  
 PN WO9707213-A2.  
 PD 27-FEB-1997.  
 PF 15-AUG-1996; U13314.  
 PR 16-AUG-1995; US-002448.

PA (HARD ) HARVARD COLLEGE.  
 PI Li J, Potter H;  
 DR WPI; 97-165297/15.  
 DR N-PSDB; T87401.  
 PT Identifying genes which cause chromosome missegregation - useful for  
 PT identifying causes of and treatments for diseases, e.g. Alzheimer's  
 PT disease, cancer and ageing  
 PS Disclosure; Fig 1; 77pp; English.  
 CC Identifying genes which cause improper chromosome segregation,  
 CC screening for inhibitors of chromosome missegregation and processes  
 CC caused by genes encoding chromosome missegregation promoters  
 CC was exemplified using Alzheimer's disease. The sequences  
 CC given in T87401 to T87426 can be used in the above methods.  
 CC It is not clear from the figure legend, the figure and the  
 CC disclosure of the specification which sequence of Fig 1 and Fig 28  
 CC is the AD4/AD3LP or the AD3 sequence.  
 SQ Sequence 376 AA;

Query Match 84.2%; Score 48; DB 1; Length 376;  
 Best Local Similarity 66.7%; Pred. No. 2.67e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 180 YPPEWS 185  
 QY 199 YPDSWS 204  
 :||:|

RESULT 4  
 ID W99829 standard; protein; 96 AA.  
 AC W99829;  
 DT 08-JUN-1999 (first entry)  
 DE HIV Vpr protein sequence #1.  
 KW HIV; Vpr; human immunodeficiency virus; hyperproliferative disease;  
 KW cell proliferation.  
 OS Human immunodeficiency virus.  
 PN WO9909412-A1.  
 PD 25-FEB-1999.  
 PF 14-AUG-1998; U16890.  
 PR 14-AUG-1997; US-055754.  
 PA (UYPE-) UNIV PENNSYLVANIA.  
 PI Ayvavoo V, Kieber-Emmons T, Mahalingam S, Patel M,  
 PI Weiner DB;  
 DR WPI; 99-181154/15.  
 PT Conjugate composition comprising HIV-1 Vpr protein fragment - used  
 PT to inhibit cell proliferation, and treating hyperproliferative  
 PT diseases  
 PS Example; Fig 1D; 64pp; English.  
 CC The present invention describes a conjugate composition comprising a  
 CC fragment of HIV-1 or non-HIV-1 Vpr protein conjugated to a therapeutic  
 CC compound. The conjugate can be used in a method for inhibiting cell  
 CC proliferation. It can also be used for treating an individual who has a  
 CC hyperproliferative disease. The HIV-1 Vpr or non-HIV-1 Vpr protein  
 CC fragments can be used for identifying compounds that inhibit Vpr protein  
 CC binding to the p6 domain of p55 or to p6 protein. The present sequence  
 CC represents an HIV protein sequence.  
 SQ Sequence 96 AA;

Query Match 82.5%; Score 47; DB 1; Length 96;  
 Best Local Similarity 66.7%; Pred. No. 3.33e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 14 YPNDWT 19  
 QY 199 YPDSWS 204  
 :||:|

RESULT 5  
 ID W99831 standard; protein; 96 AA.  
 AC W99831;  
 DT 08-JUN-1999 (first entry)  
 DE HIV E21,24P protein sequence #1.  
 KW HIV; Vpr; human immunodeficiency virus; hyperproliferative disease;  
 KW cell proliferation.

US-09-376-430-2-06.rag

Thu May 11 06:49:36 2000

OS Human immunodeficiency virus.  
 PN W990412-AL.  
 PD 23-FEB-1999.  
 PR 14-AUG-1998; U16890.  
 PR 14-AUG-1997; US-055754.  
 PR (UYPE-) UNIV PENNSYLVANIA.  
 PI Ayyavoo V, Kieber-Emmons T, Mahalingam S, Patel M,  
 PI Weiner DB;  
 DE WPI: 99-181154/15.  
 DT Conjugate composition comprising HIV-1 Vpr protein fragment - used  
 DT to inhibit cell proliferation, and treating hyperproliferative  
 DT diseases.  
 DT Example: Fig 1A: 64pp; English.  
 CC The present invention describes a conjugate composition comprising a  
 CC fragment of HIV-1 or non-HIV-1 Vpr protein conjugated to a therapeutic  
 CC compound. The conjugate can be used in a method for inhibiting cell  
 CC proliferation. It can also be used for treating an individual who has a  
 CC hyperproliferative disease. The HIV-1 Vpr or non-HIV-1 Vpr protein  
 CC fragments can be used for identifying compounds that inhibit Vpr protein  
 CC binding to the p6 domain of p55 or to p6 protein. The present sequence  
 CC represents an HIV protein sequence.  
 CC Sequence 96 AA;  
 SQ  
 Query Match 82.5%; Score 47; DB 1; Length 96;  
 Best Local Similarity 66.7%; Pred. No. 3.33e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 DB 14 YPNDWT 19  
 QY 199 YPSDWS 204  
 RESULT 6  
 ID W99812 standard; protein; 96 AA.  
 AC W99812;  
 DT 08-JUN-1999 (first entry)  
 DE HIV-1 Vpr protein.  
 DE HIV-1 Vpr; human immunodeficiency virus; hyperproliferative disease;  
 KW cell proliferation.  
 KW Human immunodeficiency virus type 1.  
 OS W09909412-AL.  
 PD 23-FEB-1999.  
 PR 14-AUG-1998; U16890.  
 PR (UYPE-) UNIV PENNSYLVANIA.  
 PI Ayyavoo V, Kieber-Emmons T, Mahalingam S, Patel M,  
 PI Weiner DB;  
 DE WPI: 99-181154/15.  
 DT Conjugate composition comprising HIV-1 Vpr protein fragment - used  
 DT to inhibit cell proliferation, and treating hyperproliferative  
 DT diseases.  
 DT Example: Fig 1A: 64pp; English.  
 CC The present invention describes a conjugate composition comprising a  
 CC fragment of HIV-1 or non-HIV-1 Vpr protein conjugated to a therapeutic  
 CC compound. The conjugate can be used in a method for inhibiting cell  
 CC proliferation. It can also be used for treating an individual who has a  
 CC hyperproliferative disease. The HIV-1 Vpr or non-HIV-1 Vpr protein  
 CC fragments can be used for identifying compounds that inhibit Vpr protein  
 CC binding to the p6 domain of p55 or to p6 protein. The present sequence  
 CC represents an HIV-1 Vpr protein sequence.  
 CC Sequence 96 AA;  
 SQ  
 Query Match 82.5%; Score 47; DB 1; Length 96;  
 Best Local Similarity 66.7%; Pred. No. 3.33e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 DB 14 YPNDWT 19  
 QY 199 YPSDWS 204  
 RESULT 7  
 ID W38504 standard; protein; 247 AA.

W38504;  
 06-NOV-1998 (first entry)  
 Streptococcus pneumoniae protein of unknown function.  
 Streptococcus pneumoniae protein; genetic immunisation; antagonist;  
 immunological response; inoculation; antibody production; inhibitor;  
 T cell immune response; antimicrobial compound; bacterial adhesion;  
 extracellular matrix protein; protein-mediated cell invasion; wound;  
 pathogenesis.  
 OS Streptococcus pneumoniae.  
 PN W09743303-AL.  
 PD 20-NOV-1997; U07950.  
 PR 14-MAY-1997; US-01670.  
 PR (SMIK) SMITHKLINE BEECHAM CORP.  
 PA (SMIK) SMITHKLINE BEECHAM PLC.  
 PI Black W, Hodgson JE, Knowles DJC, Nicholas RO,  
 PI Stodola O, 068793/01.  
 DR WPI: 98-098571.  
 DT Novel Streptococcus pneumoniae proteins and related DNA - useful for  
 DT diagnosing anti-microbial agents for treatment of bacterial  
 DT infection.  
 DT Classification: page 483pp; English.  
 CC This sequence represents a Streptococcus pneumoniae protein of  
 CC unknown function, and is encoded by a DNA sequence of the invention.  
 CC The DNA sequences were isolated from Streptococcus pneumoniae strain  
 CC 0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the  
 CC invention can be used to identify compounds which interact with and  
 CC inhibit or activate the activity of the proteins. Antagonists can be  
 CC used to treat diseases caused by S. pneumoniae proteins, through genetic  
 CC immunisation. They can also be used to induce an immunological response  
 CC in a mammal by inoculation with the S. pneumoniae proteins or delivery  
 CC of the encoding nucleic acids in a vector adequate to produce antibody  
 CC and/or T cell immune responses to protect the animal from disease. The  
 CC proteins can also be used to identify antimicrobial compounds which are  
 CC capable of inhibiting their bioactivity. In particular the proteins of  
 CC the invention can be used to prevent adhesion of bacteria to mammalian  
 CC extracellular matrix proteins on in-dwelling devices or in wounds, to  
 CC block protein-mediated mammalian cell invasion, and to block the normal  
 CC progression of pathogenesis in infections initiated other than by the  
 CC implantation of in-dwelling devices or other surgical techniques.  
 CC Note: This sequence is obtained by translating the corresponding DNA  
 CC sequence encoding this protein as the pages on which this sequence is  
 CC supposed to be was missing upon time of publication.  
 CC Sequence 247 AA;  
 SQ  
 Query Match 82.5%; Score 47; DB 1; Length 247;  
 Best Local Similarity 57.1%; Pred. No. 3.33e+02;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 DB 104 AFPADWS 110  
 QY 198 TYPDWS 204  
 RESULT 8  
 ID W05410 standard; protein; 441 AA.  
 AC W05410;  
 DT 23-FEB-1998 (first entry)  
 DE Mouse H74 protein.  
 DE Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;  
 KW cellular signalling element; cellular structural element; malignancy;  
 KW protein identification; functional domain; protein screening;  
 KW cellular signal transduction process.  
 OS Mus musculus.  
 PI Key Location/Qualifiers  
 FT Misc\_difference 439 /note= "encoded by GAC"  
 FT W09634625-AL.  
 PN 10-OCT-1996.  
 PD 04-APR-1996; U04454.  
 PR 03-APR-1996; US-630915.  
 PR 07-APR-1995; US-417872.

PA (CYTO-) CYTOGEN CORP.  
 PA (UVCN-) UNIV NORTH CAROLINA.  
 PI Fowlkes DM, Hoffman N, Kay BK, McConnell SJ, Sparks AB;  
 DR WPI: 96-465045/46.  
 DR N-PSDB: T39809.  
 DR Identifying polypeptide(s) having specific functional domain (esp.  
 PT SH3 domain) - comprises detecting selective binding to recognition  
 PT unit, regardless of sequence homology  
 PS Claim 102; Fig 47; 174pp; English.  
 CC W03405-W03411 represent human and mouse Src-homology region 3 (SH3)  
 CC domain containing proteins that can be used in the method of the  
 CC invention. SH3 domain containing proteins play a role in signalling and  
 CC structural elements of cells. The method of the invention is for  
 CC identifying polypeptides containing functional domains of interest  
 CC (especially SH3 domains). The method comprises contacting a multivalent  
 CC recognition unit (RU) complex with a number of peptides and identifying  
 CC polypeptides having a selective binding affinity for the RU complex. The  
 CC method is based on functional similarities and does not rely on sequence  
 CC similarities. Prior methods only gave limited success for identifying  
 CC proteins which contain an SH3 domain due to the minimal sequence homology  
 CC among known SH3 proteins. It has been found that small peptide RUS in  
 CC multivalent form have reduced specificity for a given functional domain  
 CC compared to monomer RUS. Multivalent RU complexes are particularly suited  
 CC to screening for polypeptides containing functional domains that are  
 CC similar to, but not identical in sequence to, the original target  
 CC functional domain. The new method enables proteins having a common  
 CC function to be identified. Identification of novel SH3 proteins will be  
 CC useful for a better understanding of cell growth, malignancy, signal  
 CC transduction processes, etc. New candidate drugs can be identified, and  
 CC their specificities (e.g. pharmacological activities) can be assessed  
 CC using the method of the invention.  
 SQ Sequence 441 AA;

Query Match 82.5%; Score 47; DB 1; Length 441;  
 Best Local Similarity 57.1%; Pred. No. 3.33e+02;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 352 TYATEWS 358  
 ||:|:|  
 QY 198 TYPDWS 204

RESULT 9  
 ID W19095 standard; Protein; 90 AA.  
 AC W19095;  
 DT 12-JAN-1998 (first entry)  
 DE Trypanosoma cruzi antigen repeat sequence.  
 KW Antigen; epitope; vaccine; protective immunity; Chagas disease;  
 KW diagnosis; therapy; immunoassay.  
 OS Trypanosoma cruzi Tulane strain C2.  
 PN W09718475-A1.  
 PD 22-MAY-1997.  
 PF 14-NOV-1996; U18624.  
 PR 14-NOV-1995; US-557309.  
 PA (CORI-) CORIXA CORP.  
 PI Houghton RL, Lodes MJ, Reed SG, Skeiky YAW;  
 DR WPI: 97-289413/26.  
 DR N-PSDB: T69153.  
 PT Diagnosing Trypanosoma cruzi infection by detecting antibodies to  
 PT novel antigens - which are useful in vaccines to provide protective  
 PT immunity against Chagas' disease  
 PS Disclosure; Page 51-52; 110pp; English.  
 CC This polypeptide sequence comprises an antigen repeat sequence  
 CC encoded by a DNA clone (see T69153) obtained by screening a  
 CC Trypanosoma cruzi genomic expression library with pools of sera  
 CC from infected individuals. Full-length T. cruzi antigens (see  
 CC W26530-41), or epitope-containing repeat sequences (see W19094-  
 CC 102, W19079-86 and W26542-44) of native antigens, can be used in a  
 CC variety of immunoassays for detecting T. cruzi infection in a  
 CC blood, serum, plasma, saliva, cerebrospinal fluid or urine sample.  
 CC The polypeptides are also useful in vaccines and pharmaceutical  
 CC compositions for inducing protective immunity against Chagas  
 CC disease. They can be produced by expression in transformed or

CC transfected host cells.  
 SQ Sequence 90 AA;

Query Match 80.7%; Score 46; DB 1; Length 90;  
 Best Local Similarity 71.4%; Pred. No. 4.15e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 25 TYPKWS 31  
 ||:|:|  
 QY 198 TYPDWS 204

RESULT 10  
 ID W42395 standard; Protein; 318 AA.  
 AC W42395;  
 DT 22-JUN-1998 (first entry)  
 DE Pyrococcus furiosus VC1 phosphatase (7ph2).  
 KW Alkaline phosphatase; thermostable enzyme; thermophilic bacterium;  
 KW food; detergent; baking.  
 OS Pyrococcus furiosus strain VC1.  
 PN W09748416-A1.  
 PD 24-DEC-1997.  
 PF 19-JUN-1997; U10784.  
 PR 19-JUN-1996; US-033752.  
 PA (RECO-) RECOMBINANT BIOCATALYSIS INC.  
 PI Bylina E, Lee E, Mathur EJ;  
 DR WPI: 98-062851/06.  
 DR N-PSDB: V03320.  
 PT Thermostable phosphatase(s) - useful in pharmaceutical, food,  
 PT detergent, and baking industries  
 PS Claim 11; Page 99-100; 128pp; English.  
 CC This protein comprises a thermostable phosphatase, designated  
 CC 7ph1, of Pyrococcus furiosus VC2. The invention relates to  
 CC claimed polynucleotides (see W42380-95). Vector and host cells  
 CC are used to produce the enzymes, which can be used in a claimed  
 CC method to hydrolyse phosphate bonds. They can also be used in  
 CC enzyme labelling processes, in certain recombinant DNA techniques,  
 CC in ELISA immunoassays, in enzyme linked gene probes, in research  
 CC applications for removing 5' phosphates in polynucleotides prior to  
 CC end labelling, and in the pharmaceutical, food, detergent, and  
 CC baking industries.  
 SQ Sequence 318 AA;

Query Match 80.7%; Score 46; DB 1; Length 318;  
 Best Local Similarity 50.0%; Pred. No. 4.15e+02;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 29 YPPEWT 34  
 ||:|:|  
 QY 199 YPDSWS 204

RESULT 11  
 ID W81970 standard; Protein; 329 AA.  
 AC W81970;  
 DT 05-FEB-1999 (first entry)  
 DE C. elegans CELF37C12.2 protein.  
 KW ER24; etoposide-induced apoptosis; degenerative disorder; p53;  
 KW cell proliferation; cell death.  
 OS Caenorhabditis elegans.  
 PN US5843659-A.  
 PD 01-DEC-1998.  
 PF 21-MAR-1996; 619362.  
 PR 21-MAR-1996; US-619362.  
 PA (APOF-) APOPTOSIS TECHNOLOGY INC.  
 PI Guild BC, Wehar SM;  
 DR WPI: 99-044568/04.  
 PT Cloned etoposide-induced apoptosis gene ER24 - useful for diagnosing  
 PT degenerative disorders characterised by inappropriate cell  
 PT proliferation or death  
 PS Example; Fig 3; 35pp; English.  
 CC This sequence represents a the Caenorhabditis elegans CELF37C12.2 protein



CC which is used to characterise a novel murine etoposide-induced apoptosis  
 CC gene, E124. The E124 gene product can be used in the diagnosis of  
 CC degenerative disorders characterised by inappropriate cell proliferation  
 CC or death. Induction of this gene by etoposide requires expression of  
 CC wild-type p53.  
 SQ Sequence 329 AA;

Query Match 80.7%; Score 46; DB 1; Length 329;  
 Best Local Similarity 42.9%; Pred. No. 4.15e+02;  
 Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 278 SYPANWN 284  
 QY 198 TYPDSWS 204

RESULT 12  
 ID R13615 standard; Protein: 337 AA.

AC R13615;  
 DT 08-NOV-1991 (first entry)  
 DE Protein found during G0 to G1 period shift.  
 KW cancer; cell proliferation.  
 OS Mus musculus.

PH Key Location/Qualifiers

FT Peptide 1..337  
 FT /label= signal peptide  
 FT peptide 27..337  
 FT /label= mature peptide  
 FT cleavage\_site 26..27  
 FT modified\_site 60  
 FT /label= N-glycosylation site  
 FT modified\_site 101  
 FT /label= N-glycosylation site  
 FT modified\_site 107  
 FT /label= N-glycosylation site  
 FT modified\_site 146  
 FT /label= N-glycosylation site  
 FT modified\_site 176  
 FT /label= N-glycosylation site  
 FT modified\_site 194  
 FT /label= N-glycosylation site  
 FT modified\_site 225  
 FT /label= N-glycosylation site  
 FT modified\_site 259  
 FT /label= N-glycosylation site  
 FT modified\_site 278  
 FT /label= N-glycosylation site

PN J03172182-A.

PD 25-JUL-1991.

PF 30-NOV-1989; 308967.

PR 30-NOV-1989; JP-308967.

PA (TOMI/) TOMINAGA S.

DR WPI; 91-263113/36.

DR N-PSDB; Q13349.

PT Use of protein coded by DNA expressed by cells - in shifting  
 PT period G0 to G1, gives information on cell proliferation useful  
 PT for anticancer drug development.

PS Claim 4; Page 2; 18pp; Japanese.

CC The protein is found in cells changing from period G0 to G1. It may  
 CC be used to investigate mechanisms of cell proliferation by using  
 CC antibodies raised to the protein. The information obtd. may be used  
 CC to develop anti cancer drugs.

SQ Sequence 337 AA;

Query Match 80.7%; Score 46; DB 1; Length 337;  
 Best Local Similarity 56.7%; Pred. No. 4.15e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0

Db 49 TYPVEW 54

QY 198 TYPDSW 203

RESULT 13

ID R99090 standard; Protein: 422 AA.

AC R99090;

DT 09-OCT-1996 (first entry)

DE Human interleukin-11 receptor.

KW Interleukin-11 receptor; IL-11 receptor; antibody; antagonist; bone;

KW osteoporosis; Paget disease; myeloma.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Peptide 1..23  
 FT /label= Sig\_peptide  
 FT protein 24..422  
 FT /label= Mat\_protein  
 FT domain 24..365  
 FT /label= Extracellular\_domain  
 FT region 24..111  
 FT /label= Ig-like\_region  
 FT region 112..365  
 FT /label= Type-1-cytokine\_region  
 FT domain 366..390  
 FT /label= Transmembrane\_domain  
 FT domain 391..422  
 FT /label= Intracellular\_domain

PN W09619574-A1.

PD 27-JUN-1996.

PF 27-NOV-1995; U15400.

PR 22-DEC-1994; US-362304.

PA (GEMY ) GENETICS INST INC.

PI Tobin JF;

DR WPI; 96-309588/31.

DR N-PSDB; R33278.

PT New nucleic acid encoding human interleukin 11 receptor - and  
 PT related protein antibodies, receptor antagonists, etc, useful for  
 PT creating and preventing loss of bone mass

PS Claim 13; Page 35-37; 54pp; English.

CC Human interleukin-11 (IL-11) receptor (R99090) is thought to play a  
 CC role in the regulation of bone maturation and repair. Its amino  
 CC acid sequence was deduced from a cDNA clone (R33278) isolated from  
 CC a human activated peripheral blood mononuclear cell cDNA library.  
 CC Recombinant IL-11 receptor or its fragments, pref. amino acids  
 CC 24-422, 24-365 (soluble extracellular domain), 391-422, 102-422 or  
 CC 102-365, can be expressed in host cell systems. It is used to  
 CC treat/prevent loss of bone mass (e.g. osteoporosis, Paget's disease,  
 CC multiple myeloma or hypogonadal conditions), as well as immune  
 CC diseases and cancer.

SQ Sequence 422 AA;

Query Match 80.7%; Score 46; DB 1; Length 422;

Best Local Similarity 57.1%; Pred. No. 4.15e+02;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 241 TYPASWP 247

QY 198 TYPDSWS 204

RESULT 14

ID R92814 standard; Protein: 423 AA.

AC R92814;

DT 21-MAY-1996 (first entry)

DE Human interleukin-11 receptor alpha chain.

KW Haemopoietin; Interleukin-11; IL-11; receptor; agonist; antagonist;

KW therapy; diagnosis.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Peptide 1..23  
 FT /label= Sig\_peptide  
 FT protein 24..423  
 FT /label= Mat\_protein  
 FT domain 24..366  
 FT /label= Extracellular\_domain  
 FT /note= "the extracellular domain includes  
 FT haemopoietin and 19-like domains"

Search completed: Wed May 10 11:53:03 2000  
Job time : 7 secs.

```
FT domain 367, 392
FT /label= Transmembrane_domain
FT 393, 423
FT domain
FT /label= Cytoplasmic_tail
PN WO9607737-A1.
PD 14-MAR-1996.
PF 05-SEP-1995; AU0578.
PR 05-SEP-1994; AU-007902.
PR 05-SEP-1994; AU-007901.
PA (AMRA-) AMRAD OPERATIONS PTY LTD.
PI Hilton DJ;
DR WPI; 96-171612/17.
DR N-PSDB; T17869.
PT Nucleic acid encoding haemopoietin receptor containing conserved
PT amino acid motif esp. IL-11 receptor alpha chain - used for
PT developing IL-11 (ant)agonists
PS Claim 8; Page 47-49; 87pp; English.
CC The human interleukin-11 (IL-11) receptor alpha chain (R92814)
CC was identified by expression of DNA (T17869) isolated from human
CC bone marrow cDNA libraries. Expression of the human IL-11
CC receptor alpha chain results in specific binding of human IL-11
CC and permits IL-11 signalling. The receptor alpha chain can be used
CC to develop agonists or antagonists of therapeutic appln. or in
CC the treatment or diagnosis of conditions involving a deficiency of
CC IL-11, excess IL-11 or aberrant effects of normal endogenous IL-11
CC levels.
SQ Sequence 423 AA;

Query Match 80.7%; Score 46; DB 1; Length 423;
Best Local Similarity 57.1%; Pred. No. 4.15e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 242 TYPASWP 248
QY 198 TYPSDWS 204
||||:|

RESULT 15
ID R06260 standard; protein; 517 AA.
AC R06260.
DT 07-DEC-1990 (first entry)
DE Human acetylcholine receptor TE671 (Achr) delta-subunit.
KW Nicotinic acetyl choline receptor; Achr; TE671; insecticides;
KW Muscle relaxants; anthelmintics.
OS Homo sapiens.
PN CA2003459-A.
PD 23-MAY-1990.
PF 21-NOV-1989; 003459.
PR 23-NOV-1988; US-275422.
PA (SALK ) SALK INST FOR BIOL STUD.
PI Lindstrom JM, Schoepfer RD;
DR WPI; 90-231525/31.
DR N-PSDB; Q05558.
PT Human muscle nicotinic acetyl:choline receptor - used to assay
PT the effects of agents which affect acetyl:choline receptors in
PT skeletal muscles.
PS Disclosure; English.
CC Receptors may be used in assay for materials which modify them.
CC They may be produced in substantial, pure quantities for use in
CC experimentation, development of insecticides without effect on
CC hMNARS and treatment of parasitic infections. Mabs raised to the
CC peptides may be useful in detection of the structure of MNARS.
SQ Sequence 517 AA;

Query Match 80.7%; Score 46; DB 1; Length 517;
Best Local Similarity 66.7%; Pred. No. 4.15e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 192 TYPVEW 197
QY 198 TYPSDW 203
||||:|
```

\*\*\*\*\*  
 M P S R L H (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Mpsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 11:53:21 2000; MasPar time 44.49 Seconds  
 Tabular output not generated. 2.039 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (198-204) from US09376430A.pap (6 of 25)  
 Perfect Score: 57  
 Sequence: 1 TYPEDSWS 7

Scoring table: PAM 150  
 Gap 11

Searched: 131253 seqs, 12956647 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: a-issued  
 1:5A\_COMB 2:5B\_COMB 3:PCT\_COMB 4:backfiles1

Statistics: Mean 15.758; Variance 56.096; scale 0.281

pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	48	84.2	376	2	US-08-875-Sequence 2, Applicatio	1.65e+02
2	46	80.7	390	2	US-08-552-Sequence 24, Applicati	2.54e+02
3	46	80.7	329	2	US-08-615-Sequence 10, Applicati	2.54e+02
4	46	80.7	337	1	US-08-442-Sequence 18, Applicati	2.54e+02
5	45	78.9	199	2	US-08-665-Sequence 9, Applicatio	3.14e+02
6	45	78.9	273	2	US-08-424-Sequence 9, Applicatio	3.14e+02
7	45	78.9	273	2	US-08-820-Sequence 9, Applicatio	3.14e+02
8	45	78.9	273	2	US-08-826-Sequence 9, Applicatio	3.14e+02
9	45	78.9	278	3	PCT-US94-0Sequence 2, Applicatio	3.14e+02
10	45	78.9	279	2	US-08-474-Sequence 30, Applicati	3.14e+02
11	45	78.9	279	3	PCT-US91-0Sequence 29, Applicati	3.14e+02
12	45	78.9	279	1	US-07-688-Sequence 30, Applicati	3.14e+02
13	45	78.9	872	1	US-08-451-Sequence 8, Applicatio	3.87e+02
14	44	77.2	137	2	US-07-857-Sequence 110, Applicat	3.87e+02
15	44	77.2	467	1	US-08-140-Sequence 2, Applicatio	4.77e+02
16	43	75.4	261	1	US-08-175-Sequence 2, Applicatio	4.77e+02
17	43	75.4	261	1	US-07-971-Sequence 4, Applicatio	4.77e+02
18	43	75.4	588	3	PCT-US95-1Sequence 2, Applicatio	4.77e+02
19	43	75.4	616	1	US-08-385-Sequence 4, Applicatio	4.77e+02
20	43	75.4	616	1	US-08-385-Sequence 2, Applicatio	4.77e+02
21	43	75.4	2431	1	US-07-920-Sequence 2, Applicatio	4.77e+02
22	43	75.4	2627	1	US-09-060-Sequence 3, Applicatio	4.77e+02
23	43	75.4	2627	2	US-08-751-Sequence 3, Applicatio	4.77e+02

ALIGNMENTS

RESULT ID	1	US-08-875-972-2	STANDARD:	PRT:	376 AA.
XX	AC	xxxxxx			
XX	DT				
XX	XX	Sequence 2, Application US/08875972			
DE	XX	Sequence 2, Application US/08875972			
XX	CC	Patent No. 595554			
CC	CC	GENERAL INFORMATION:			
CC	CC	APPLICANT: Huntington Potter and Jinhue Li			
CC	CC	TITLE OF INVENTION: ASSAY FOR IDENTIFYING GENES CAUSING			
CC	CC	TITLE OF INVENTION: CHROMOSOME NON-DISJUNCTION			
CC	CC	NUMBER OF SEQUENCES: 29			
CC	CC	CORRESPONDENCE ADDRESS:			
CC	CC	ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.			
CC	CC	STREET: Two Militia Drive			
CC	CC	CITY: Lexington			
CC	CC	STATE: Massachusetts			
CC	CC	COUNTRY: USA			
CC	CC	ZIP: 02173-4799			
CC	CC	COMPUTER READABLE FORM:			
CC	CC	MEDIUM TYPE: Floppy disk			
CC	CC	OPERATING SYSTEM: IBM PC compatible			
CC	CC	SOFTWARE: Patent In Release #1.0, Version #1.30			
CC	CC	CURRENT APPLICATION DATA:			
CC	CC	APPLICATION NUMBER: US/08/875,972			
CC	CC	FILING DATE: 08-AUG-97			
CC	CC	CLASSIFICATION: 435			
CC	CC	PRIOR APPLICATION DATA:			
CC	CC	APPLICATION NUMBER: US 60/002,448			
CC	CC	FILING DATE: 16-AUG-1995			
CC	CC	ATTORNEY/AGENT INFORMATION:			
CC	CC	NAME: Granahan Esq., Patricia			
CC	CC	REGISTRATION NUMBER: 32,227			
CC	CC	REFERENCE/DOCKET NUMBER: HU95-03PA			
CC	CC	TELECOMMUNICATION INFORMATION:			
CC	CC	TELEPHONE: (781) 861-6240			
CC	CC	TELEFAX: (781) 861-9540			
CC	CC	INFORMATION FOR SEQ ID NO: 2:			
CC	CC	SEQUENCE CHARACTERISTICS:			
CC	CC	LENGTH: 376 amino acids			
CC	CC	TYPE: amino acid			
CC	CC	TOPOLOGY: linear			

24	42	73.7	21	2	US-08-463-Sequence 310, Applicati	5.87e+02
25	42	73.7	21	1	US-08-465-Sequence 14, Applicati	5.87e+02
26	42	73.7	84	2	US-08-353-Sequence 94, Applicati	5.87e+02
27	42	73.7	85	1	US-08-451-Sequence 54, Applicati	5.87e+02
28	42	73.7	85	1	US-08-450-Sequence 54, Applicati	5.87e+02
29	42	73.7	85	1	US-08-450-Sequence 54, Applicati	5.87e+02
30	42	73.7	85	1	US-08-455-Sequence 23, Applicati	5.87e+02
31	42	73.7	121	1	US-08-450-Sequence 51, Applicati	5.87e+02
32	42	73.7	121	1	US-08-450-Sequence 51, Applicati	5.87e+02
33	42	73.7	121	1	US-08-451-Sequence 51, Applicati	5.87e+02
34	42	73.7	121	1	US-08-450-Sequence 38, Applicati	5.87e+02
35	42	73.7	134	1	US-08-450-Sequence 38, Applicati	5.87e+02
36	42	73.7	143	1	US-08-450-Sequence 63, Applicati	5.87e+02
37	42	73.7	143	1	US-08-450-Sequence 63, Applicati	5.87e+02
38	42	73.7	148	1	US-08-450-Sequence 59, Applicati	5.87e+02
39	42	73.7	148	1	US-08-450-Sequence 59, Applicati	5.87e+02
40	42	73.7	148	1	US-08-450-Sequence 60, Applicati	5.87e+02
41	42	73.7	157	1	US-08-450-Sequence 69, Applicati	5.87e+02
42	42	73.7	240	1	US-08-461-Sequence 2, Applicatio	5.87e+02
43	42	73.7	314	1	US-08-859-Sequence 6, Applicatio	5.87e+02
44	42	73.7	335	2	US-08-459-Sequence 2, Applicatio	5.87e+02
45	42	73.7	993	1	US-08-468-Sequence 2, Applicatio	5.87e+02

```

CC MOLECULE TYPE: protein
SQ SEQUENCE 376 AA; 42057 MW; 833360 CN;

Query Match      84.2%; Score 48; DB 2; Length 376;
Best Local Similarity 66.7%; Pred.No. 1.65e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 180 YPPEWS 185
   ||:|||
QY 199 YPSDWS 204

RESULT 2
ID US-08-557-309B-24 STANDARD; PRT; 90 AA.
XX XXXXXX
XX
XX
XX DT
XX
DE XX
XX
CC Sequence 24, Application US/08557309B
CC Sequence 24, Application US/08557309B
CC Patent No. 5916572
CC GENERAL INFORMATION:
CC APPLICANT: Reed, Steven G.
CC APPLICANT: Skeiky, Yasir A.W.
CC APPLICANT: Lodes, Michael J.
CC APPLICANT: Houghton, Raymond L.
CC TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION
CC NUMBER OF SEQUENCES: 69 CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SEED and BERRY LLP
CC CITY: Seattle
CC STATE: Washington
CC COUNTRY: USA
CC ZIP: 98104-7092
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/557,309B
CC FILING DATE: 14-NOV-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Maki, David J.
CC REGISTRATION NUMBER: 31,392
CC REFERENCE/DOCKET NUMBER: 210121.422
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 622-4900
CC TELEFAX: (206) 682-6031
CC INFORMATION FOR SEQ ID NO: 24:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 90 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC
SQ SEQUENCE 90 AA; 10757 MW; 45517 CN;

Query Match      80.7%; Score 46; DB 2; Length 90;
Best Local Similarity 71.4%; Pred.No. 2.54e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Ddb 25 TYPKAWS 31
   |||||
QY 198 TYPSDWS 204

RESULT 3
ID US-08-619-362A-10 STANDARD; PRT; 329 AA.

```

AC	xxxxxx
XX	
DT	
XX	
DE	
XX	
CC	Sequence 10, Application US/08619362A
CC	Sequence 10, Application US/08619362A
CC	Patent No. 5843659
CC	GENERAL INFORMATION:
CC	APPLICANT: SOPHIE M. LEHAR; and
CC	APPLICANT: BRAYDON C. GUILD
CC	TITLE OF INVENTION: NOVEL APOPTOSIS GENE E124,
CC	TITLE OF INVENTION: COMPOSITIONS, AND METHODS OF USE
CC	NUMBER OF SEQUENCES: 17
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: Hale and Dorr
CC	STREET: 1455 Pennsylvania Avenue, N.W.
CC	CITY: Washington
CC	STATE: District of Columbia
CC	COUNTRY: U.S.A.
CC	ZIP: 20004
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: Floppy disk
CC	COMPUTER: IBM PC compatible
CC	OPERATING SYSTEM: PC-DOS/MS-DOS
CC	SOFTWARE: PatentIn Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: US/08/619,362A
CC	FILING DATE: 21 MARCH 1996
CC	CLASSIFICATION: 435
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: WIXON, Henry N.
CC	REGISTRATION NUMBER: 32,073
CC	REFERENCE/DOCKET NUMBER: 104322.170
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: (202) 942-8400
CC	TELEFAX: (202) 942-8484
CC	INFORMATION FOR SEQ ID NO: 10:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 329 amino acids
CC	TYPE: amino acid
CC	STRANDEDNESS: single
CC	TOPOLOGY: linear
SQ	SEQUENCE 329 AA; 38080 MW; 596419 CN;
	Query Match 80.7%; Score 46; DB 2; Length 329;
	Best Local Similarity 42.9%; Pred.No. 2.54e+02;
	Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Db	278 SYPANWN 284
QY	:  :: :
	198 TYPDWS 204
RESULT	4
ID	US-08-442-043A-18 STANDARD; PRT; 337 AA.
XX	xxxxxx
AC	
XX	
DT	
XX	
DE	
XX	
CC	Sequence 18, Application US/08442043A
CC	Sequence 18, Application US/08442043A
CC	Patent No. 5767064
CC	GENERAL INFORMATION:
CC	APPLICANT: Sims, John E.
CC	APPLICANT: Cosman, David J.
CC	APPLICANT: Lupton, Stephen D.
CC	APPLICANT: Mosley, Bruce A.
CC	APPLICANT: Dower, Steven K.
CC	TITLE OF INVENTION: Type II Interleukin-1 Receptors
CC	NUMBER OF SEQUENCES: 19

US-08-376-430-2-06.ra1

Thu May 11 06:49:37 2000

CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Immunex Corporation  
 CC STREET: 51 University Street  
 CC CITY: Seattle  
 CC STATE: WA  
 CC COUNTRY: USA  
 CC ZIP: 98101  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC Compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent Release #1.24  
 CC CURRENT APPLICATION DATA:  
 CC FILING DATE: 16-MAY-1995  
 CC APPLICATION NUMBER: US/08/442,043A  
 CC FILING DATE: 16-MAY-1995  
 CC CLASSIFICATION: 514  
 CC PRIOR APPLICATION DATA:  
 CC PRIOR APPLICATION NUMBER: US 07/534,193  
 CC FILING DATE: 05-JUN-1990  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07/573,576  
 CC FILING DATE: 24-AUG-1990  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07/627,071  
 CC FILING DATE: 13-DEC-1990  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07/701,415  
 CC FILING DATE: 16-MAY-1991  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/091,519  
 CC FILING DATE: 12-JULY-1993  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/242,211  
 CC FILING DATE: 13-MAY-1994  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Perkins, Patricia A.  
 CC REGISTRATION NUMBER: 34,693  
 CC REFERENCE/DOCKET NUMBER: 2003-F  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 206-587-5570  
 CC TELEFAX: 206-233-0644  
 CC INFORMATION FOR SEQ ID NO: 18:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 337 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE:  
 CC ORGANISM: Murine  
 CC IMMEDIATE SOURCE:  
 CC CLONE: MUST2  
 CC SEQUENCE 337 AA; 38501 MW; 582048 CN;

Query Match 80.7%; Score 46; DB 1; Length 337;  
 Best Local Similarity 66.7%; Pred No. 2.54e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 49 TYPVEW 54  
 QY 198 TYPEDW 203

RESULT 5  
 ID US-08-665-647-9 STANDARD; PRT; 199 AA.

XX XXXXXX

Sequence 9, Application US/08665647  
 Sequence 9, Application US/08665647

CC Patent No. 5935803  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Dasquez, Nicki J.  
 CC APPLICANT: Ron, Dorit  
 CC APPLICANT: Voronova, Anna F.  
 CC APPLICANT: Napolitano, Eugene W.  
 CC TITLE OF INVENTION: METHODS TO IDENTIFY IMMUNOMODULATORS  
 CC TITLE OF INVENTION: USING COGNATE INTERACTION OF PKC-THETA  
 CC NUMBER OF SEQUENCES: 89  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: MORRISON & FOERSTER  
 CC STREET: 2000 Pennsylvania Avenue, NW - Ste. 5500  
 CC CITY: Washington  
 CC STATE: DC  
 CC COUNTRY: USA  
 CC ZIP: 20006-1888  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC Compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/665,647  
 CC FILING DATE: 18-JUN-1996  
 CC CLASSIFICATION: 435  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Murashige, Kate H.  
 CC REGISTRATION NUMBER: 29,959  
 CC REFERENCE/DOCKET NUMBER: 22550-20025.25  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (202) 887-1500  
 CC TELEFAX: (202) 822-0168  
 CC TELEX: 90-4030 MRSNFOERSWSH  
 CC INFORMATION FOR SEQ ID NO: 9:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 199 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 199 AA; 22434 MW; 210125 CN;

Query Match 78.9%; Score 45; DB 2; Length 199;  
 Best Local Similarity 66.7%; Pred. No. 3.14e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 8 TYPENW 13  
 QY 198 TYPEDW 203

RESULT 6  
 ID US-08-424-641B-9 STANDARD; PRT; 273 AA.

XX XXXXXX

Sequence 9, Application US/08424641B  
 Sequence 9, Application US/08424641B  
 Patent No. 5824523  
 GENERAL INFORMATION:  
 APPLICANT: Sylvain Moineau, Shirley A.  
 APPLICANT: Walker, Ebenezer R. Vedamuthu,  
 APPLICANT: and Peter A. Vandenbergh  
 TITLE OF INVENTION: Isolated DNA Encoding  
 TITLE OF INVENTION: Enzyme For Phage  
 TITLE OF INVENTION: Resistance  
 NUMBER OF SEQUENCES: 12  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Ian C. McLeod  
 STREET: 2190 Commons Parkway  
 CITY: Okemos

CC STATE: Michigan  
CC COUNTRY: USA  
CC ZIP: 48864  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb  
CC MEDIUM TYPE: storage  
CC COMPUTER: Acer  
CC OPERATING SYSTEM: MS-DOS  
CC SOFTWARE: Wordperfect 5.1  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/424,641B  
CC FILING DATE: April 19, 1995  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 08/366,480  
CC FILING DATE: December 30, 1994  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Ian C. McLeod  
CC REGISTRATION NUMBER: 20,931  
CC REFERENCE/DOCKET NUMBER: MT 4.1-151  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (517) 347-4100  
CC TELEFAX: (517) 347-4103  
CC TELEX: No. 5824523e  
CC INFORMATION FOR SEQ ID NO: 9:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 273 Amino Acids  
CC TYPE: Amino Acid  
CC STRANDEDNESS: Single  
CC TOPOLOGY: Linear  
CC MOLECULE TYPE: Peptide  
CC SEQUENCE 273 AA; 31923 MW; 388074 CN;

Query Match 78.9%; Score 45; DB 2; Length 273;  
Best Local Similarity 66.7%; Pred. No. 3.14e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 138 TYSSEW 143  
||:|:  
QY 198 TYPSPW 203

RESULT 7  
ID US-08-820-980-9 STANDARD; PRT; 273 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE  
Sequence 9, Application US/08820980  
XX  
Sequence 9, Application US/08820980  
CC Patent No. 5925388  
CC GENERAL INFORMATION:  
CC APPLICANT: Sylvain Moineau, Shirley A.  
CC APPLICANT: Walker, Ebenezer R. Vedamuthu,  
CC APPLICANT: and Peter A. Vandenberg  
CC TITLE OF INVENTION: Isolated DNA Encoding  
CC TITLE OF INVENTION: Enzyme For Phage  
CC TITLE OF INVENTION: Resistance  
CC NUMBER OF SEQUENCES: 12  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Ian C. McLeod  
CC STREET: 2190 Commons Parkway  
CC CITY: Okemos  
CC STATE: Michigan  
CC COUNTRY: USA  
CC ZIP: 48864  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb  
CC MEDIUM TYPE: storage  
CC COMPUTER: Acer

CC OPERATING SYSTEM: MS-DOS  
CC SOFTWARE: Wordperfect 5.1  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/820,980  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 08/424,641  
CC FILING DATE: April 19, 1995  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Ian C. McLeod  
CC REGISTRATION NUMBER: 20,931  
CC REFERENCE/DOCKET NUMBER: Quest 4.1-156  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (517) 347-4100  
CC TELEFAX: (517) 347-4103  
CC TELEX: No. 5925388e  
CC INFORMATION FOR SEQ ID NO: 9:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 273 Amino Acids  
CC TYPE: Amino Acid  
CC STRANDEDNESS: Single  
CC TOPOLOGY: Linear  
CC MOLECULE TYPE: Peptide  
CC SEQUENCE 273 AA; 31923 MW; 388074 CN;

Query Match 78.9%; Score 45; DB 2; Length 273;  
Best Local Similarity 66.7%; Pred. No. 3.14e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 138 TYSSEW 143  
||:|:  
QY 198 TYPSPW 203

RESULT 8  
ID US-08-826-439-9 STANDARD; PRT; 273 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE  
Sequence 9, Application US/08826439  
XX  
Sequence 9, Application US/08826439  
CC Patent No. 5972673  
CC GENERAL INFORMATION:  
CC APPLICANT: Sylvain Moineau, Shirley A.  
CC APPLICANT: Walker, Ebenezer R. Vedamuthu,  
CC APPLICANT: and Peter A. Vandenberg  
CC TITLE OF INVENTION: Isolated DNA Encoding  
CC TITLE OF INVENTION: Enzyme For Phage  
CC TITLE OF INVENTION: Resistance  
CC NUMBER OF SEQUENCES: 12  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Ian C. McLeod  
CC STREET: 2190 Commons Parkway  
CC CITY: Okemos  
CC STATE: Michigan  
CC COUNTRY: USA  
CC ZIP: 48864  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb  
CC MEDIUM TYPE: storage  
CC COMPUTER: Acer  
CC OPERATING SYSTEM: MS-DOS  
CC SOFTWARE: Wordperfect 5.1  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/826,439  
CC FILING DATE:  
CC CLASSIFICATION: 530  
CC PRIOR APPLICATION DATA:

US-09-376-430-2-06.rai

Thu May 11 06:49:37 2000

CC APPLICATION NUMBER: 08/424,641  
 CC FILING DATE: April 19, 1995  
 CC CLASSIFICATION: 530  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Ian C. McLeod  
 CC REGISTRATION NUMBER: 20,931  
 CC REFERENCE/DOCKET NUMBER: Quest 4.1-155  
 CC TELEPHONE: (517) 347-4100  
 CC TELEFAX: (517) 347-4103  
 CC TELEX: NO. 5972673e  
 CC INFORMATION FOR SEQ ID NO: 9:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 273 Amino Acids  
 CC TYPE: Amino Acid  
 CC STRANDEDNESS: Single  
 CC TOPOLOGY: Linear  
 CC MOLECULE TYPE: Peptide  
 CC SEQUENCE 273 AA; 31923 MW; 388074 CN;  
 CC  
 CC Query Match 78.9%; Score 45; DB 2; Length 273;  
 CC Best Local Similarity 66.7%; Pred. No. 3.14e+02;  
 CC Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 Db 138 TYSSEM 143  
 QY 198 TYPDSW 203  
 CC  
 CC RESULT 9  
 CC ID PCT-US94-03744-2 STANDARD; PRT; 278 AA.  
 CC XX  
 CC AC xxxxxx  
 CC XX  
 CC DT  
 CC DE  
 CC SEQUENCE 2, Application PC/TUS9403744  
 CC GENERAL INFORMATION:  
 CC APPLICANT: PETRI, WILLIAM A.  
 CC APPLICANT: MCCOY, JAMES J.  
 CC APPLICANT: MANN, BARBARA J.  
 CC TITLE OF INVENTION: 35/31 KDA SUBUNIT OF THE  
 CC TITLE OF INVENTION: ENTAMOEBA  
 CC TITLE OF INVENTION: HISTOLYTICA ADHERENCE LECTIN  
 CC NUMBER OF SEQUENCES: 11  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: MORRISON & Foerster  
 CC STREET: 2000 Pennsylvania Avenue, Suite 5500  
 CC CITY: Washington  
 CC STATE: DC  
 CC COUNTRY: USA  
 CC ZIP: 20006-1812  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent In Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: PCT/US94/03744  
 CC FILING DATE:  
 CC CLASSIFICATION:  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/045,679  
 CC FILING DATE: 09-APR-1993  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: MURASHIGE, KATE H.  
 CC REGISTRATION NUMBER: 29,959  
 CC REFERENCE/DOCKET NUMBER: 29148-20005.00  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (202) 887-1500  
 CC TELEFAX: (202) 887-0763

CC TELEX: 90-4030  
 CC INFORMATION FOR SEQ ID NO: 2:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 278 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 278 AA; 32461 MW; 409664 CN;  
 CC  
 CC Query Match 78.9%; Score 45; DB 3; Length 278;  
 CC Best Local Similarity 60.0%; Pred. No. 3.14e+02;  
 CC Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 Db 186 YPTBW 190  
 QY 199 YPSDW 203  
 CC  
 CC RESULT 10  
 CC ID US-08-474-379C-30 STANDARD; PRT; 279 AA.  
 CC XX  
 CC AC xxxxxx  
 CC XX  
 CC DT  
 CC DE  
 CC SEQUENCE 30, Application US/08474379C  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Wigler, Michael H.  
 CC APPLICANT: Colicelli, John J.  
 CC TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED  
 CC TITLE OF INVENTION: PROCESSES  
 CC NUMBER OF SEQUENCES: 88  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 CC STREET: 233 South Wacker Drive/6300 Sears Tower  
 CC CITY: Chicago  
 CC STATE: Illinois  
 CC COUNTRY: United States of America  
 CC ZIP: 60606-6402  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent In Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/474,379C  
 CC FILING DATE: 07-JUN-1995  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07/511,715  
 CC FILING DATE: 20-APR-1990  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/206,188  
 CC FILING DATE: 01-MAR-1994  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07/688,352  
 CC FILING DATE: 19-APR-1991  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Clough, David W.  
 CC REGISTRATION NUMBER: 36,107  
 CC REFERENCE/DOCKET NUMBER: 27866/32771  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (312) 474-6300  
 CC TELEFAX: (312) 474-0448  
 CC INFORMATION FOR SEQ ID NO: 30:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 279 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein

SQ SEQUENCE 279 AA; 31850 MW; 413959 CN;  
Query Match 78.9%; Score 45; DB 2; Length 279;  
Best Local Similarity 71.4%; Pred. No. 3.14e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 190 AYPDWS 196  
:| | | |  
QY 198 TYPDWS 204

RESULT 11  
ID PCT-US91-02714-29 STANDARD; PRT; 279 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE  
XX  
Sequence 29, Application PC/TUS9102714  
CC  
Sequence 29, Application PC/TUS9102714  
CC  
GENERAL INFORMATION:  
CC  
APPLICANT: Wigler, Michael H.  
CC  
APPLICANT: Colicelli, John J.  
CC  
TITLE OF INVENTION: Cloning by Complementation and Related  
CC  
TITLE OF INVENTION: Processes  
CC  
NUMBER OF SEQUENCES: 55  
CC  
CORRESPONDENCE ADDRESS:  
CC  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
CC  
ADDRESSEE: Bicknell  
CC  
STREET: Two First National Plaza, 20 South Clark  
CC  
STREET: Street  
CC  
CITY: Chicago  
CC  
STATE: Illinois  
CC  
COUNTRY: USA  
CC  
ZIP: 60603  
CC  
COMPUTER READABLE FORM:  
CC  
MEDIUM TYPE: Floppy disk  
CC  
COMPUTER: IBM PC compatible  
CC  
OPERATING SYSTEM: PC-DOS/MS-DOS  
CC  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CC  
CURRENT APPLICATION DATA:  
CC  
APPLICATION NUMBER: PCT/US91/02714  
CC  
FILING DATE: 19910419  
CC  
CLASSIFICATION: 435  
CC  
PRIOR APPLICATION DATA:  
CC  
APPLICATION NUMBER: US 07/511,715  
CC  
FILING DATE: 20-APR-1990  
CC  
ATTORNEY/AGENT INFORMATION:  
CC  
NAME: Borun, Michael F.  
CC  
REGISTRATION NUMBER: 25447  
CC  
REFERENCE/DOCKET NUMBER: 27805/30197  
CC  
TELEPHONE: (312) 346-5750  
CC  
TELEFAX: (312) 984-9740  
CC  
TELEX: 25-3856  
CC  
INFORMATION FOR SEQ ID NO: 29:  
CC  
SEQUENCE CHARACTERISTICS:  
CC  
LENGTH: 279 amino acids  
CC  
TYPE: AMINO ACID  
CC  
TOPOLOGY: linear  
CC  
MOLECULE TYPE: protein  
CC  
SEQUENCE 279 AA; 31850 MW; 413959 CN;

Query Match 78.9%; Score 45; DB 3; Length 279;  
Best Local Similarity 71.4%; Pred. No. 3.14e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 190 AYPDWS 196  
:| | | |  
QY 198 TYPDWS 204

RESULT 13  
ID US-08-451-715A-8 STANDARD; PRT; 872 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE  
XX  
Sequence 8, Application US/08451715A  
CC  
Sequence 8, Application US/08451715A  
CC

SQ SEQUENCE 279 AA; 31850 MW; 413959 CN;  
Query Match 78.9%; Score 45; DB 1; Length 279;  
Best Local Similarity 71.4%; Pred. No. 3.14e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 190 AYPDWS 196  
:| | | |  
QY 198 TYPDWS 204

RESULT 12  
ID US-07-688-352C-30 STANDARD; PRT; 279 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE  
XX  
Sequence 30, Application US/07688352C  
CC  
Sequence 30, Application US/07688352C  
CC  
Patent No. 5527896  
CC  
GENERAL INFORMATION:  
CC  
APPLICANT: Wigler, Michael H.  
CC  
APPLICANT: Colicelli, John J.  
CC  
TITLE OF INVENTION: Cloning by Complementation and Related  
CC  
TITLE OF INVENTION: Processes  
CC  
NUMBER OF SEQUENCES: 57  
CC  
CORRESPONDENCE ADDRESS:  
CC  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
CC  
ADDRESSEE: Bicknell  
CC  
STREET: Two First National Plaza, 20 South Clark  
CC  
STREET: Street  
CC  
CITY: Chicago  
CC  
STATE: Illinois  
CC  
COUNTRY: USA  
CC  
ZIP: 60603  
CC  
COMPUTER READABLE FORM:  
CC  
MEDIUM TYPE: Floppy disk  
CC  
COMPUTER: IBM PC compatible  
CC  
OPERATING SYSTEM: PC-DOS/MS-DOS  
CC  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CC  
CURRENT APPLICATION DATA:  
CC  
APPLICATION NUMBER: US/07/688,352C  
CC  
FILING DATE: 19910419  
CC  
CLASSIFICATION: 435  
CC  
PRIOR APPLICATION DATA:  
CC  
APPLICATION NUMBER: US 07/511,715  
CC  
FILING DATE: 20-APR-1990  
CC  
ATTORNEY/AGENT INFORMATION:  
CC  
NAME: Borun, Michael F.  
CC  
REGISTRATION NUMBER: 25447  
CC  
REFERENCE/DOCKET NUMBER: 27805/30197  
CC  
TELEPHONE: (312) 346-5750  
CC  
TELEFAX: (312) 984-9740  
CC  
TELEX: 25-3856  
CC  
INFORMATION FOR SEQ ID NO: 30:  
CC  
SEQUENCE CHARACTERISTICS:  
CC  
LENGTH: 279 amino acids  
CC  
TYPE: AMINO ACID  
CC  
TOPOLOGY: linear  
CC  
MOLECULE TYPE: protein  
CC  
SEQUENCE 279 AA; 31850 MW; 413959 CN;

Query Match 78.9%; Score 45; DB 1; Length 279;  
Best Local Similarity 71.4%; Pred. No. 3.14e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 190 AYPDWS 196  
:| | | |  
QY 198 TYPDWS 204



Thu May 11 06:49:37 2000

CC Patent No. 5801013  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Tao, Jianshi  
 CC APPLICANT: Qui, Yan  
 CC APPLICANT: Houman, Fariba  
 CC APPLICANT: Shen, Xiaoyu  
 CC APPLICANT: Schimmel, Paul R.  
 CC TITLE OF INVENTION: Helicobacter Aminoacyl-tRNA Synthetase  
 CC TITLE OF INVENTION: Proteins, Nucleic Acids and Strains Comprising Same  
 CC NUMBER OF SEQUENCES: 67  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 CC STREET: Two Militia Drive  
 CC CITY: Lexington  
 CC STATE: Massachusetts  
 CC COUNTRY: USA  
 CC ZIP: 02173  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/451,715A  
 CC FILING DATE: 26-MAY-1995  
 CC CLASSIFICATION: 435  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Brook, David E.  
 CC REGISTRATION NUMBER: 22,592  
 CC REFERENCE/DOCKET NUMBER: CP194-25  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 617-861-6240  
 CC TELEFAX: 617-861-9540  
 CC INFORMATION FOR SEQ ID NO: 8:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 872 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 872 AA; 101171 MW; 3891778 CN;  
 CC  
 CC Query Match 78.9%; Score 45; DB 1; Length 872;  
 CC Best Local Similarity 80.0%; Pred. No. 3.14e+02;  
 CC Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC Db 385 YPSNW 389  
 CC QY 199 YPSDW 203  
 CC  
 CC RESULT 14  
 CC ID US-07-857-224B-110 STANDARD; PRT; 137 AA.  
 CC XX  
 CC AC xxxxxx  
 CC XX  
 CC DT  
 CC DE  
 CC XX  
 CC Sequence 110, Application US/07857224B  
 CC  
 CC Sequence 110, Application US/07857224B  
 CC Patent No. 5958784  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Benner, Steven A.  
 CC TITLE OF INVENTION: Predicting Folded Structures of Proteins  
 CC NUMBER OF SEQUENCES: 114  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Steven A. Benner  
 CC STREET: Hadlaubstrasse 151  
 CC CITY: Zurich  
 CC STATE: none  
 CC COUNTRY: Switzerland  
 CC ZIP: (note: this is an international post code) CH-8092  
 CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage  
 CC COMPUTER: Apple Macintosh  
 CC OPERATING SYSTEM: Macintosh 7.0  
 CC SOFTWARE: Microsoft Word  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/07/857,224B  
 CC FILING DATE: 03/25/92  
 CC CLASSIFICATION: 436  
 CC PRIOR APPLICATION DATA: none  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (International) 41 1 632 2830  
 CC TELEFAX: (International) 41 1 262 2437  
 CC TELEX: none  
 CC INFORMATION FOR SEQ ID NO: 110:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 137  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC DESCRIPTION: protein  
 CC ORIGINAL SOURCE:  
 CC ORGANISM: rat  
 CC FEATURE: sperm coating glycoprotein; Table 17 Row 6  
 CC PUBLICATION INFORMATION:  
 CC AUTHORS:  
 CC AUTHORS: Brooks, D. E.  
 CC AUTHORS: Means, A. R.  
 CC AUTHORS: Wright, E. J.  
 CC AUTHORS: Singh, S. P.  
 CC AUTHORS: Tiver, K. K.  
 CC TITLE: Molecular cloning of the cDNA for androgen-dependent sperm  
 CC TITLE: coating glycoprotein secreted by the rat epidemis.  
 CC JOURNAL: European Journal of Biochemistry  
 CC VOLUME: 161  
 CC PAGES: 13-18  
 CC DATE: 1986  
 CC SEQUENCE 137 AA; 15622 MW; 107827 CN;  
 CC  
 CC Query Match 77.2%; Score 44; DB 2; Length 137;  
 CC Best Local Similarity 66.7%; Pred. No. 3.87e+02;  
 CC Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC Db 45 YPSAWS 50  
 CC QY 199 YPSDWS 204  
 CC  
 CC RESULT 15  
 CC ID US-08-140-104A-2 STANDARD; PRT; 467 AA.  
 CC XX  
 CC AC xxxxxx  
 CC XX  
 CC DT  
 CC DE  
 CC XX  
 CC Sequence 2, Application US/08140104A  
 CC  
 CC Sequence 2, Application US/08140104A  
 CC Patent No. 5585255  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Tsukada, Yogi  
 CC APPLICANT: Tazuke, Yasuhiko  
 CC APPLICANT: Okada, Shigenori  
 CC APPLICANT: Adachi, Kenichi  
 CC TITLE OF INVENTION: BILE ACID SULFATASE GENE, PLASMD  
 CC TITLE OF INVENTION: CONTAINING SAID GENE AND METHOD OF PRODUCING BILE ACID  
 CC TITLE OF INVENTION: SULFATE SULFATASE  
 CC NUMBER OF SEQUENCES: 4  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Knobbe, Martens, Olson and Bear  
 CC STREET: 620 Newport Center Drive 16th Floor  
 CC CITY: Newport Beach  
 CC STATE: CA  
 CC COUNTRY: USA

CC ZIP: 92660  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patentin Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/140,104A  
 CC FILING DATE:  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: JP PCT/JP93/00244  
 CC FILING DATE: 26-FEB-1993  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Altman, Daniel E.  
 CC REGISTRATION NUMBER: 34,115  
 CC REFERENCE/DOCKET NUMBER: SAEGU3.001AUS  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 714-760-0404  
 CC TELEFAX: 714-760-9502  
 CC INFORMATION FOR SEQ ID NO: 2:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 467 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 467 AA; 51349 MW; 1038998 CN;

Query Match 77.2%; Score 44; DB 1; Length 467;  
 Best Local Similarity 50.0%; Pred. No. 3.87e+02;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 251 YPEWA 256  
 QY 199 YPSDWS 204

Search completed: Wed May 10 11:54:16 2000  
 Job time : 55 secs.

Thu May 11 06:49:37 2000

\*\*\*\*\*  
Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd.  
\*\*\*\*\*  
MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:52:31 2000; MasPar time 3.87 Seconds  
Tabular output not generated.  
Title: >US-09-376-430-2  
Description: (198-204) from US09376430A.pep (6 of 25)  
Perfect score: 57  
Sequence: 1 TYPEDWS 7  
Scoring table: PAM 150  
Gap 11  
Searched: 142080 seqs, 47172406 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: pir62  
1:pir1 2:pir2 3:pir3 4:pir4  
Statistics: Mean 22.700; Variance 31.747; scale 0.715  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.  
SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	51	89.5	261	2	T05447	8.60e+00
2	49	86.0	112	2	T08219	2.01e+01
3	48	84.2	166	2	F69029	3.04e-01
4	48	84.2	230	2	T00557	3.04e-01
5	48	84.2	419	2	S76571	3.04e-01
6	48	84.2	476	2	S71360	3.04e-01
7	48	84.2	953	2	S19427	3.04e-01
8	47	82.5	218	2	I39802	4.57e+01
9	47	82.5	287	2	A72751	4.57e+01
10	47	82.5	621	1	S59632	4.57e+01
11	47	82.5	850	1	WMBE56	4.57e+01
12	46	80.7	102	2	D71042	6.84e+01
13	46	80.7	266	2	J00864	6.84e+01
14	46	80.7	337	2	G4842	6.84e+01
15	46	80.7	375	2	C38530	6.84e+01
16	46	80.7	422	2	T37891	6.84e+01
17	46	80.7	435	2	I07047	6.84e+01
18	46	80.7	436	2	S20060	6.84e+01
19	46	80.7	436	2	I51237	6.84e+01
20	46	80.7	437	2	I51238	6.84e+01
21	46	80.7	517	2	A60916	6.84e+01
22	46	80.7	517	2	S29498	6.84e+01
23	46	80.7	567	2	S29498	6.84e+01

24	46	80.7	592	2	D70863	hypothetical protein	6.84e+01
25	46	80.7	1036	2	H64245	hypothetical protein	6.84e+01
26	46	80.7	1072	1	SYBY14	isoleucine--trna liga	6.84e+01
27	46	80.7	1782	2	S45289	vitellogenin precursor	6.84e+01
28	46	80.7	4436	2	E71086	hypothetical protein	6.84e+01
29	45	78.9	51	2	S36376	ig heavy chain v regi	1.02e+02
30	45	78.9	59	2	S29578	ig light chain - rain	1.02e+02
31	45	78.9	140	2	S04575	ig heavy chain precu	1.02e+02
32	45	78.9	167	2	S29579	ig light chain - rain	1.02e+02
33	45	78.9	232	2	G34284	H+-transporting ATP s	1.02e+02
34	45	78.9	235	2	S20000	ig light chain precu	1.02e+02
35	45	78.9	273	1	S35647	site-specific DNA met	1.02e+02
36	45	78.9	278	2	A49505	lectin light chain -	1.02e+02
37	45	78.9	431	2	I76773	ATP binding component	1.02e+02
38	45	78.9	432	2	I48343	interleukin-11 recept	1.02e+02
39	45	78.9	437	1	S25649	translation elongatio	1.02e+02
40	45	78.9	471	2	B71400	glycine hydroxymethyl	1.02e+02
41	45	78.9	787	2	A64522	virB4 homolog - Helic	1.02e+02
42	45	78.9	855	2	A53296	DNA mismatch repair p	1.02e+02
43	45	78.9	868	21	D63297	transmembrane oligosa	1.02e+02
44	45	78.9	874	2	A64664	valine--trna ligase (	1.02e+02
45	45	78.9	3623	2	T08618	intrinsic factor-B12	1.02e+02

ALIGNMENTS

RESULT 1  
ENTRY T05447 #type complete  
TITLE hypothetical protein F7K2.110 - Arabidopsis thaliana  
ORGANISM #formal\_name Arabidopsis thaliana #common\_name mouse-ear  
cress  
DATE 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change  
23-Jul-1999  
ACCESSIONS T05447  
REFERENCE Z15416  
#authors Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes,  
H.W.; Mayer, K.F.X.; Schueller, C.  
#submission submitted to the Protein Sequence Database, November 1998  
#accession T05447  
#molecule\_type DNA  
##residues 1-261 #label BEV  
##cross-references EMBL:AL03545  
##experimental\_source cultivar Columbia; BAC clone F7K2  
GENETICS  
#map\_position 4  
#note 50/3  
#introns F7K2.110  
#trans F7K2.110  
SUMMARY #length 261 #molecular-weight 29106 #checksum 5763  
Query Match 89.5%; Score 51; DB 2; Length 261;  
Best Local Similarity 83.3%; Pred. No. 8.60e+00;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Db 18 TYPADW 23  
Qy 198 TYPADW 203  
RESULT 2  
ENTRY T08219 #type complete  
TITLE hypothetical protein H0011 - Halobacterium sp. (strain NRC-1)  
ORGANISM #formal\_name Halobacterium sp.  
#variety strain NRC-1  
DATE 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change  
T08219  
ACCESSIONS Z16408  
REFERENCE Ng, W.V.; Clufo, S.A.; Smith, T.M.; Bumgarner, R.E.; Baskin,  
D.; Faust, J.; Hall, B.; Loretz, C.; Seto, J.; Slagel, J.;  
Hood, L.; Dassarma, S.  
#authors Genome Res. (1998) 8:1131-1141  
#journal Snapshot of a large dynamic replicon in a halophilic  
#title

```

Archaeon: megaplasmid or minichromosome?
#cross-references MUID:99063795
#accession T08219
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 1-112 ##label NGW
##cross-references EMBL:AF016485; NID:g2822278; PID:g2822280
##experimental_source strain NRC-1
GENETICS
#genome plasmid pNRC100
SUMMARY
#length 112 #molecular-weight 12807 #checksum 5632
Query Match 86.0%; Score 49; DB 2; Length 112;
Best Local Similarity 71.4%; Pred. No. 2.01e+01;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 35 TYPGWA 41
QY 198 TYPGWS 204
RESULT 3
ENTRY F69029 #type complete
TITLE hypothetical protein MTH1218 - Methanobacterium
ORGANISM thermoautotrophicum (strain Delta H)
DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
ACCESSIONS F69029
REFERENCE A69000
#authors Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.;
Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.;
Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.;
Lumm, W.; Pothier, B.; Qiu, D.; Spadafora, R.; Vicaire, K.;
Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwanji, N.; Caruso,
A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.;
McDougall, S.; Shimer, G.; Goyal, A.; Pietrokovski, S.;
Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling,
J.; Reeve, J.N.
#journal J. Bacteriol. (1997) 179:7135-7155
#title Complete genome sequence of Methanobacterium
thermoautotrophicum Delta H: functional analysis and
comparative genomics.
#cross-references MUID:98037514
#accession F69029
##status preliminary; nucleic acid sequence not shown;
translation not shown
##molecule_type DNA
##residues 1-166 ##label MTH
##cross-references GB:AE000889; GB:AE000666; NID:g2622318; PID:g2622328
##experimental_source strain Delta H
GENETICS
#gene MTH1218
SUMMARY
#length 166 #molecular-weight 18277 #checksum 1591
Query Match 84.2%; Score 48; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 3.04e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 33 YPSDW 37
QY 199 YPSDW 203
RESULT 4
ENTRY T00557 #type complete
TITLE hypothetical protein F12L6.13 - Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
cress
DATE 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change
ACCESSIONS T00557
REFERENCE Z14168

```

```

#authors Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.;
Brandon, R.C.; Sykes, S.M.; Kaul, S.; Mason, T.M.;
Kerlavage, A.R.; Adams, M.D.; Somerville, C.R.; Venter,
J.C.
#submission submitted to the EMBL Data Library, July 1998
#description Arabidopsis thaliana chromosome II BAC F12L6 genomic
sequence.
#accession T00557
##status translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 1-230 ##label ROU
##cross-references EMBL:AC004218; NID:g3355463; PID:g3355476
##experimental_source cultivar Columbia
GENETICS
#map_position 2
#introns 20/3; 40/1; 75/1; 99/3; 140/3; 164/3; 197/1
#note F12L6.13
SUMMARY
#length 230 #molecular-weight 26171 #checksum 7577
Query Match 84.2%; Score 48; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 3.04e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 94 YPSDW 98
QY 199 YPSDW 203
RESULT 5
ENTRY S76571 #type complete
TITLE hypothetical protein - Synecocystis sp. (strain PCC 6803)
ORGANISM #formal_name Synecocystis sp.
#variety PCC 6803
DATE 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change
ACCESSIONS S76571
REFERENCE S74322
#authors Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.;
Nakamura, Y.; Miyajima, N.; Hirose, M.; Sugita, M.;
Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.;
Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimpo,
S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.;
Yasuda, M.; Tabata, S.
#journal DNA Res. (1996) 3:109-136
#title Sequence analysis of the genome of the unicellular
Cyanobacterium Synecocystis sp. PCC6803. II. Sequence
determination of the entire genome and assignment of
potential protein-coding regions.
#cross-references MUID:97061201
#accession S76571
##status preliminary
##molecule_type DNA
##residues 1-419 ##label KAN
##cross-references EMBL:D64002; GB:AB001339; NID:g1001612; PID:d1011068;
PID:g1001682
##note the nucleotide sequence was submitted to the EMBL Data
Library, June 1996
SUMMARY
#length 419 #molecular-weight 46994 #checksum 6335
Query Match 84.2%; Score 48; DB 2; Length 419;
Best Local Similarity 100.0%; Pred. No. 3.04e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 206 YPSDW 210
QY 199 YPSDW 203
RESULT 6
ENTRY S71360 #type complete
TITLE noEA protein - Rhizobium meliloti
ORGANISM #formal_name Rhizobium meliloti
DATE 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change

```

613-629 #domain transmembrane #status predicted #label TM14\  
643-659 #domain transmembrane #status predicted #label TM15  
#length 953 #molecular-weight 107897 #checksum 7258  
SUMMARY  
Query Match 84.2%; Score 48; DB 2; Length 953;  
Best Local Similarity 66.7%; Pred. No. 3.04e+01;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Db 43 TYPDEW 48  
QY 198 TYPSDW 203  
RESULT 8  
ENTRY #type complete  
TITLE motB protein homolog - Bacillus megaterium  
ORGANISM #formal\_name Bacillus megaterium  
DATE 19-Jul-1996 #sequence\_revision 19-Jul-1996 #text\_change 13-Nov-1998  
ACCESSIONS I39802  
REFERENCE Hueck, C.; Kraus, A.; Hillen, W.D.R.  
#authors Gene (1994) 143:147-148  
#journal Sequences of ccpA and two downstream Bacillus megaterium  
#title genes with homology to the motAB operon from Bacillus subtilis.  
#cross-references MUID:94259294  
#accession I39802  
#status preliminary; translated from GB/EMBL/DDBJ  
#molecule\_type DNA  
#residues 1-218 #label RES  
#cross-references GB:L26052; NID:g415663; PID:g415666  
GENETICS  
#start\_codon GTC  
#length 218 #molecular-weight 24642 #checksum 8705  
SUMMARY  
Query Match 82.5%; Score 47; DB 2; Length 218;  
Best Local Similarity 66.7%; Pred. No. 4.57e+01;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Db 152 AVPSNW 157  
QY 198 TYPSDW 203  
RESULT 9  
ENTRY #type complete  
TITLE hypothetical protein APE0534 - Aeropyrum pernix (strain K1)  
ORGANISM #formal\_name Aeropyrum pernix  
DATE 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
ACCESSIONS A72751  
REFERENCE A72450  
#authors Kawaiabavasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.;  
Haikawa, Y.; Jin-no, K.; Takahashi, M.; Sekine, M.; Baba, S.;  
Ankai, A.; Kosugi, H.; Hosoyama, A.; Fukui, S.; Nagai, Y.;  
Nishijima, K.; Nakazawa, H.; Takamiya, M.; Masuda, S.;  
Funahashi, I.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;  
Kushida, N.; Oguchi, A.; Aoki, K.; Kubota, K.; Nakamura, Y.;  
Nomura, N.; Sako, Y.; Kikuchi, H.  
#journal DNA Res. (1999) 6:83-101  
#title Complete genome sequence of an aerobic hyper-thermophilic  
#cross-references MUID:99310339  
#accession A72751  
#status preliminary  
#molecule\_type DNA  
#residues 1-287 #label KAW  
#cross-references DDBJ:AP000059; NID:g5103911; PID:BAA79501.1;  
#experimental\_source strain K1  
GENETICS  
#gene APE0534

06-Feb-1998  
S71360; S71358  
REFERENCE  
#authors Ardourel, M.; Lortet, G.; Maillet, F.; Roche, P.; Truchet, G.;  
Prome, J.C.; Rosenberg, C.  
#submission submitted to the EMBL Data Library, May 1995  
#accession S71360  
#molecule\_type DNA  
#residues 1-476 #label ARD  
#cross-references EMBL:U26430; NID:g1326070  
#experimental\_source strain RCR2011  
REFERENCE S71357  
#authors Ardourel, M.; Lortet, G.; Maillet, F.; Roche, P.; Truchet, G.;  
Prome, J.C.; Rosenberg, C.  
#journal Mol. Microbiol. (1995) 17:687-699  
#title In Rhizobium meliloti, the operon associated with the nod box  
n5 comprises nodL, noeA and noeB, three host-range genes  
specifically required for the nodulation of particular  
Medicago species.  
#cross-references MUID:96111489  
#accession S71358  
#status nucleic acid sequence not shown  
#molecule\_type DNA  
#residues 304-339 #label ARW  
#cross-references EMBL:U26430  
#experimental\_source strain RCR2011 (-SU47)  
GENETICS  
#gene noeA  
#keywords nodulation  
SUMMARY  
#length 476 #molecular-weight 53673 #checksum 3841  
Query Match 84.2%; Score 48; DB 2; Length 476;  
Best Local Similarity 71.4%; Pred. No. 3.04e+01;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Db 96 TYPYEW 102  
QY 198 TYPSDWS 204  
RESULT 7  
ENTRY #type complete  
TITLE probable membrane protein YCR017c - yeast (Saccharomyces cerevisiae)  
ORGANISM #formal\_name Saccharomyces cerevisiae  
DATE 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 30-Jan-1998  
ACCESSIONS S19427  
REFERENCE S19424  
#authors Hatat, D.; Jacq, C.; Perea, J.; Shu, Y.  
#submission submitted to the Protein Sequence Database, March 1992  
#accession S19427  
#molecule\_type DNA  
#residues 1-953 #label HAT  
#cross-references EMBL:X59720; NID:g1907116; PID:e264481; PID:g1907160  
MIPS:YCR017c  
GENETICS  
#map\_position 3R  
FEATURES  
8-24 transmembrane protein  
8-24 #domain transmembrane #status predicted #label TM1\  
65-81 #domain transmembrane #status predicted #label TM2\  
92-108 #domain transmembrane #status predicted #label TM3\  
121-137 #domain transmembrane #status predicted #label TM4\  
308-324 #domain transmembrane #status predicted #label TM5\  
332-348 #domain transmembrane #status predicted #label TM6\  
353-369 #domain transmembrane #status predicted #label TM7\  
381-397 #domain transmembrane #status predicted #label TM8\  
423-439 #domain transmembrane #status predicted #label TM9\  
455-471 #domain transmembrane #status predicted #label TM10\  
522-538 #domain transmembrane #status predicted #label TM11\  
548-564 #domain transmembrane #status predicted #label TM12\  
575-591 #domain transmembrane #status predicted #label TM13\  
575-591

```

SUMMARY          #length 287 #molecular-weight 28767 #checksum 5732
Query Match      82.5%; Score 47; DB 2; Length 287;
Best Local Similarity 80.0%; Pred. No. 4.57e+01;
Matches          4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 24 YPSEW 28
|||||
QY 199 YPSDW 203

RESULT 10
ENTRY  #molecule_type DNA
TITLE  endo-1,4-beta-xylanase (EC 3.2.1.18) B precursor - Cellvibrio
        mixtus
ORGANISM #formal_name Cellvibrio mixtus
DATE 15-Feb-1996 #sequence_revision 11-Apr-1997 #text_change
ACCESSIONS S59632; S52742
REFERENCE S59631
#authors Millward-Sadler, S.J.; Davidson, K.; Hazlewood, G.P.; Black,
#journal G.W.; Gilbert, H.J.; Clarke, J.H.
#title Novel cellulose-binding domains, NodB homologues and
        conserved modular architecture in xylanases from the
        aerobic soil bacteria Pseudomonas fluorescens subsp.
        cellulosa and Cellvibrio mixtus.
#cross-references MIMD:96077124
#accession S59632
#molecule_type DNA
#residues 1-621 #label MIL
#cross-references EMBL:248926; NID:g757808; PIDN:CAA88762.1;
        PID:g757809

GENETICS
#gene xynB
FUNCTION #description catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in
        xylans
#pathway xylan degradation
CLASSIFICATION #superfamily Pseudomonas endo-1,4-beta-xylanase F;
        Streptomyces endo-1,4-beta-xylanase A homology
KEYWORDS glycosylase; hydrolase; polysaccharide degradation
FEATURE 1-19
20-621 #domain signal sequence #status predicted #label SIG\
        #product endo-1,4-beta-xylanase B #status predicted
        #label MAT\
302-615 #domain Streptomyces endo-1,4-beta-xylanase A homology
        #label SXI\
403,516 #active_site Glu #status predicted
SUMMARY #length 621 #molecular-weight 64929 #checksum 1375

Query Match      82.5%; Score 47; DB 1; Length 621;
Best Local Similarity 42.9%; Pred. No. 4.57e+01;
Matches          3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 183 YPGEWG 189
|||||
QY 198 TYPDWS 204

RESULT 11
ENTRY  #type complete
TITLE  infected cell protein ICP18.5 - human cytomegalovirus (strain
        AD169)
ALTERNATE_NAMES HFLF0 protein
ORGANISM #formal_name human cytomegalovirus, human herpesvirus 5
#note host Homo sapiens (man)
DATE 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change
ACCESSIONS S09819
REFERENCE S09749
#authors Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.;
        Cerny, R.; Horsnell, T.; Hutchinson III, C.A.; Kouzarides,
        T.; Martignetti, J.A.; Preddie, E.; Satchwell, S.C.;
        Tomlinson, P.; Weston, K.M.; Barrell, B.G.
        Curr. Top. Microbiol. Immunol. (1990) 154:125-169
        Analysis of the protein-coding content of the sequence of
        human cytomegalovirus strain AD169.
#cross-references MIMD:90269039
#accession S09819
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-850 #label CHE
#cross-references EMBL:X17403; NID:g59591; PIDN:CAA35371.1;
        PID:g1780834
#note possible protein-coding frames are given
#note the DNA sequence was submitted to EMBL, December 1989,
        in computer-readable form
CLASSIFICATION #superfamily herpesvirus infected cell protein ICP18.5
KEYWORDS capsid assembly
SUMMARY #length 850 #molecular-weight 95867 #checksum 6125

Query Match      82.5%; Score 47; DB 1; Length 850;
Best Local Similarity 80.0%; Pred. No. 4.57e+01;
Matches          4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 667 YPSEW 671
|||||
QY 199 YPSDW 203

RESULT 12
ENTRY  #type complete
TITLE  hypothetical protein PH1628 - Pyrococcus horikoshii
ORGANISM #formal_name Pyrococcus horikoshii
DATE 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change
        14-Aug-1998
ACCESSIONS D71042
REFERENCE A71000
#authors Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.;
        Hino, Y.; Yamamoto, S.; Sekine, M.; Baba, S.; Kosugi, H.;
        Hosoyama, A.; Nagai, Y.; Sakai, M.; Ogura, K.; Otsuka, R.;
        Nakazawa, H.; Takamiya, M.; Ohfuku, Y.; Funahashi, T.;
        Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,
        A.; Aoki, K.; Yoshizawa, T.; Nakamura, Y.; Robb, F.T.;
        Horikoshi, K.; Masuchi, Y.; Shizuya, H.; Kikuchi, H.
#journal DNA Res. (1998) 5:55-76
#title Complete sequence and gene organization of the genome of a
        hyper-thermophilic archaeobacterium, Pyrococcus horikoshii
        OT3.
#cross-references MIMD:98344137
#accession D71042
#status preliminary; nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
#residues 1-102 #label KAW
#cross-references GB:AP000006; NID:g3236133; PID:d1031683; PID:g3258057
#experimental_source strain OT3
#note this accession replaces an interim accession for a
        sequence replaced by GenBank

GENETICS
#gene PH1628
SUMMARY #length 102 #molecular-weight 10865 #checksum 7199

Query Match      80.7%; Score 46; DB 2; Length 102;
Best Local Similarity 57.1%; Pred. No. 6.84e+01;
Matches          4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 40 SYPSAWT 46
|||||
QY 198 TYPDWS 204

RESULT 13
ENTRY  #type complete
TITLE  hypothetical 21.9K protein - Escherichia coli retron Ec67

```

US-09-376-430-2-06.rpr

Thu May 11 06:49:37 2000

```

ALTERNATE_NAMES
ORGANISM      ci protein homolog
DATE          12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change
10-Sep-1997

ACCESSIONS
REFERENCE      JQ0864
#authors      Hsu, M.Y.; Inouye, M.; Inouye, S.
#journal      Proc. Natl. Acad. Sci. U.S.A. (1990) 87:9454-9458
#title        Retron for the 67-base multicopy single-stranded DNA from
              Escherichia coli: a potential transposable element encoding
              both reverse transcriptase and Dam methylase functions.
#cross-references MIM:91067724
#accession     JQ0864
#molecule_type DNA
#residues      1-198 #label HSU
#cross-references GB:M55249; NID:g145143; PID:g145144
#experimental_source E. coli strain Cl-1
GENETICS
#note          Insertion site is equivalent to 19 min of E. coli K12 genetic
              map
SUMMARY
#length 198 #molecular-weight 21898 #checksum 7751
Query Match      80.7%; Score 46; DB 2; Length 198;
Best Local Similarity 56.7%; Pred. No. 6.84e+01;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 56 TTPADW 61
|:|:|
QY 198 YPDSW 203

RESULT 14
ENTRY      G64842 #type complete
#authors    Blattnher, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.;
              Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.;
              Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.;
              Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao
              Y.
#journal     Science (1997) 277:1453-1462
#title       The complete genome sequence of Escherichia coli K-12.
#cross-references MIM:97426617
#accession   G64842
#molecule_type DNA
#residues    1-266 #label BLAT
#cross-references GB:AE000202; GB:U00096; NID:g1787233; PID:g1787244;
              UWGP:bl1009
#experimental_source strain K-12, substrain MG1655
#length 266 #molecular-weight 28898 #checksum 5702
SUMMARY
Query Match      80.7%; Score 46; DB 2; Length 266;
Best Local Similarity 80.0%; Pred. No. 6.84e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 149 YPADW 153
|:|:|
QY 199 YPDSW 203

RESULT 15
ENTRY      A33541 #type complete
#authors    ST2 protein precursor - mouse
#journal     38.5K T1 glycoprotein
#title       ALTERNATE_NAMES #formal_name Mus musculus #common_name house mouse
              ORGANISM      21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change
              DATE          10-Sep-1997
              ACCESSIONS      A33541; S17657; S07054

```

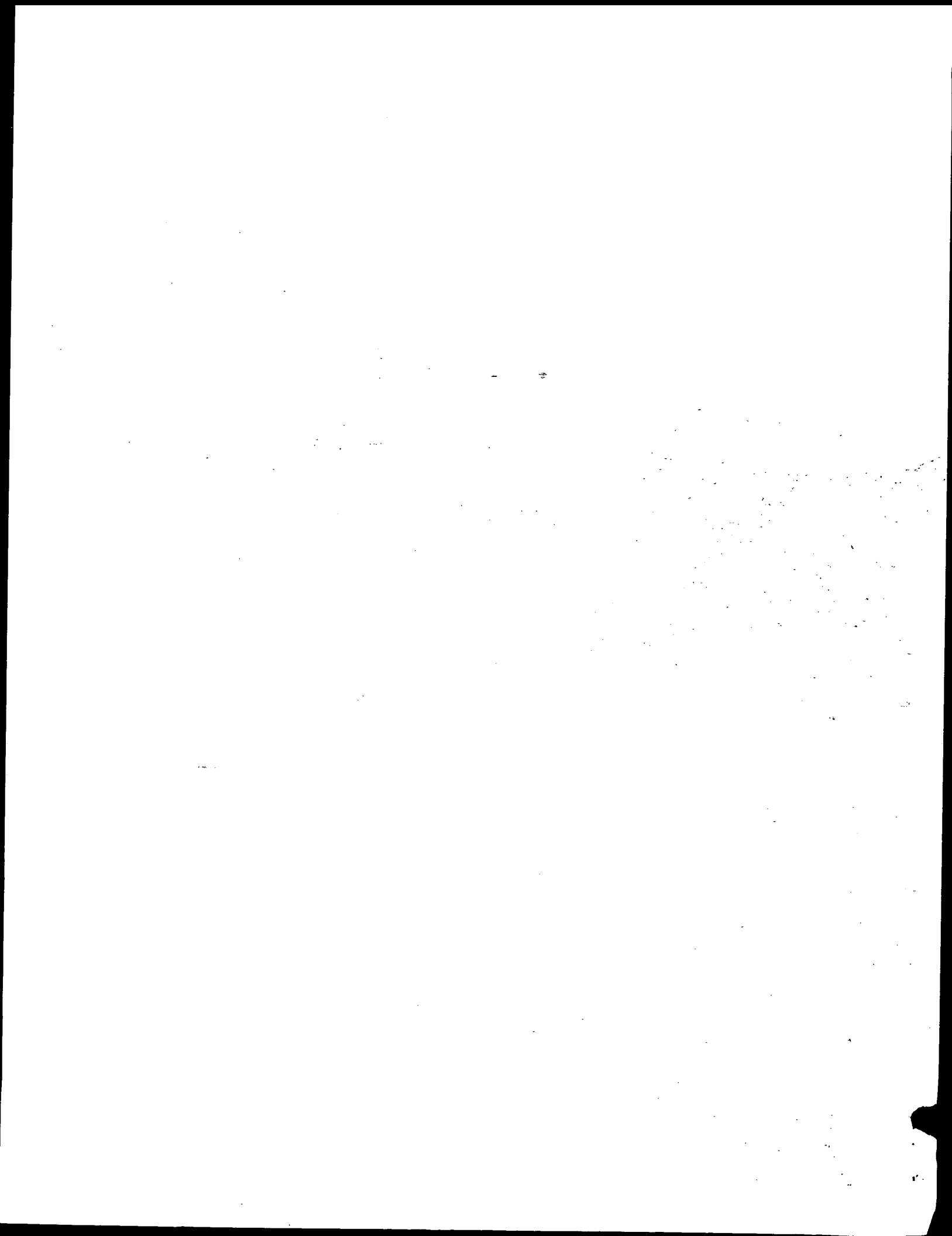
```

REFERENCE      A33541
#authors      Klemenzt, R.; Hoffmann, S.; Wersniskold, A.K.
#journal      Proc. Natl. Acad. Sci. U.S.A. (1989) 86:5708-5712
#title        Serum- and oncoprotein-mediated induction of a gene with
              sequence similarity to the gene encoding carcinoembryonic
              antigen.
#cross-references MIM:89345536
#accession     A33541
#molecule_type mRNA
#residues      1-337 #label KLE
#cross-references GB:M24843; NID:g201103; PID:g201104
              S17657
#authors      Tomimaga, S.I.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.;
              Tetsuka, T.
#journal      Biochim. Biophys. Acta (1991) 1090:1-8
#title        Molecular cloning of the murine ST2 gene. Characterization
              and chromosomal mapping.
#cross-references MIM:91355215
#accession     S17657
#status        translation not shown
#molecule_type DNA
#residues      1-337 #label TOM
#cross-references EMBL:X60184; NID:g54200; PID:g54201
              S07054
#authors      Tomimaga, S.I.
#journal      FEBS Lett. (1989) 258:301-304
#title        A putative protein of a growth specific cDNA from BALB/c-3T3
              cells is highly similar to the extracellular portion of
              mouse interleukin 1 receptor.
#cross-references MIM:90092495
#accession     S07054
#molecule_type mRNA
#residues      1-191, A' 193-337 #label TO2
#cross-references EMBL:Y07519; NID:g55517; PID:g55518
#note          it is uncertain whether Met-1, Met-7 or Met-19 is the
              initiator
GENETICS
#gene         ST2
#map_position 1
#introns      27/1; 97/2; 155/3; 210/1; 233/1; 280/2
KEYWORDS      glycoprotein
FEATURE
1-26          #domain signal sequence #status predicted #label SIG\
27-337        #product ST2 protein #status predicted #label MAT\
60,101,107,145,176, #binding_site carbohydrate (Asn) (covalent) #status
194,225,259,278      predicted
SUMMARY        #length 337 #molecular-weight 38530 #checksum 5054
Query Match      80.7%; Score 46; DB 2; Length 337;
Best Local Similarity 66.7%; Pred. No. 6.84e+01;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 49 TYPVEM 54
|:|:|
QY 198 TYPVEM 203

Search completed: Wed May 10 11:52:40 2000
Job time : 9 secs.

```





\*\*\*\*\*  
 M P E R L H  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 11:46:20 2000; Maspar time 90.00 seconds  
 Tabular output not generated. 2.369 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (198-204) from US09376430A.pep (6 of 25)  
 Perfect Score: 57  
 Sequence: 1 TYPEDWS 7

Scoring table: PAM 150  
 Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: swiss-prot38  
 1:swissprot

Statistics: Mean 23.191; Variance 29.414; scale 0.786

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	48	84.2	953	1	YCQ7_YEAST	1.28e+01
2	47	82.5	218	1	YTXE_BACME	1.99e+01
3	47	82.5	250	1	ECHR_ERWCH	1.99e+01
4	47	82.5	367	1	CARA_SULSO	1.99e+01
5	47	82.5	850	1	PROBABLE_PROCESSING	1.99e+01
6	46	80.7	198	1	YR7I_ECOLI	3.06e+01
7	46	80.7	337	1	ST2_MOUSE	3.06e+01
8	46	80.7	375	1	TGT_SHIFT	3.06e+01
9	46	80.7	375	1	TGT_SHIFT	3.06e+01
10	46	80.7	436	1	ELONGATION_FACTOR_1-GA	3.06e+01
11	46	80.7	436	1	ELONGATION_FACTOR_1-GA	3.06e+01
12	46	80.7	517	1	ACHD_HUMAN	3.06e+01
13	46	80.7	1036	1	ACETYLCHOLINE_RECEPTOR	3.06e+01
14	46	80.7	1072	1	Y414_MYGE	3.06e+01
15	46	80.7	1782	1	SYIC_YEAST	3.06e+01
16	46	80.7	3093	1	VIT_BOMMO	3.06e+01
17	45	78.9	178	1	POLG_BSTV1	3.06e+01
18	45	78.9	232	1	YCB7_HUMAN	3.06e+01
19	45	78.9	242	1	ATP6_PARLI	3.06e+01
20	45	78.9	249	1	YTXE_BACSU	3.06e+01
21	45	78.9	250	1	ESAR_ERWST	3.06e+01
22	45	78.9	273	1	Y564_HUMAN	3.06e+01
23	45	78.9	273	1	MTIC_MORBO	3.06e+01

24	45	78.9	288	1	GIL2_ENTHI	GALACTOSE-INHIBITABLE	4.69e+01
25	45	78.9	292	1	YAY4_YEAST	HYPOTHETICAL 34.5 KD P	4.69e+01
26	45	78.9	437	1	EF1G_RABIT	ELONGATION FACTOR 1-GA	4.69e+01
27	45	78.9	437	1	EF1G_HUMAN	ELONGATION FACTOR 1-GA	4.69e+01
28	45	78.9	798	1	PHSG_BACSU	GLYCOCEN PHOSPHORILASE	4.69e+01
29	45	78.9	855	1	MUTS_AZOVI	DNA MISMATCH REPAIR PR	4.69e+01
30	45	78.9	874	1	SVY_HELPY	VALYL-TRNA SYNTHETASE	4.69e+01
31	45	78.9	1064	1	SYIC_SCHPO	ISOLEUCYL-TRNA SYNTHET	4.69e+01
32	45	78.9	4540	1	DYHC_PARTIE	DYNEIN HEAVY CHAIN, CY	4.69e+01
33	44	77.2	175	1	YAFM_HABCH	HYPOTHETICAL 28.4 KD P	7.13e+01
34	44	77.2	254	1	YVLI_VIBCH	HYPOTHETICAL 38.6 KD P	7.13e+01
35	44	77.2	354	1	YQJL_ECOLI	HYPOTHETICAL 52.6 KD P	7.13e+01
36	44	77.2	465	1	YQJX_HABIN	HYPOTHETICAL PROTEIN H	7.13e+01
37	44	77.2	470	1	YQJX_HABIN	HYPOTHETICAL PROTEIN H	7.13e+01
38	44	77.2	692	1	VNUC_WABVP	NUCLEOPROTEIN (NUCLEOC	7.13e+01
39	44	77.2	695	1	VNUC_WABVP	NUCLEOPROTEIN (NUCLEOC	7.13e+01
40	44	77.2	697	1	BYN_DROME	T-RELATED PROTEIN (TRP	7.13e+01
41	44	77.2	909	1	HEX_ADEMI	HEXON PROTEIN (LAME PR	7.13e+01
42	44	77.2	921	1	SYI_BACSU	ISOLEUCYL-TRNA SYNTHET	7.13e+01
43	44	77.2	986	1	YD43_SCHPO	HYPOTHETICAL 111.1 KD	7.13e+01
44	44	77.2	1331	1	Y064_MYCGE	HYPOTHETICAL PROTEIN M	7.13e+01
45	44	77.2	1386	1	Y064_MYCPN	HYPOTHETICAL PROTEIN M	7.13e+01

## ALIGNMENTS

RESULT	1	STANDARD;	PRT;	953 AA.
ID	YCQ7_YEAST			
AC	P25618;			
DT	01-MAY-1992 (Rel. 22, Created)			
DT	01-MAY-1992 (Rel. 22, Last sequence update)			
DE	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	HYPOTHETICAL 107.9 KD PROTEIN IN POLA-SRDI INTERGENIC REGION.			
GN	YCQ017C OR YCR17C.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;			
OC	Saccharomycetales; Saccharomycetes.			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RA	Hatata D., Jacq C., Perea J., Shu Y.;			
RL	Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; X59720; CAA42308.1;			
DR	PIR; S19427; S19427.			
KW	Hypothetical protein.			
SQ	SEQUENCE 953 AA; 107897 MW; 01FEL7EFC3BFE15 CRC64;			
Query Match	84.2%; Score 48; DB 1; Length 953;			
Best local similarity	66.7%; Pred. No. 1.28e+01;			
Matches	4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
Db	43 TYPEDW 48			
Qy	198 TYPEDW 203			
RESULT	2	STANDARD;	PRT;	218 AA.
ID	YTXE_BACME			
AC	P46827;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	HYPOTHETICAL 24.6 KD PROTEIN IN CCPA 3 REGION (ORF2).			
GN	YTXE.			
OS	Bacillus megaterium.			

Bacteria; Firmicutes; Bacillus/Clostridium group;  
Bacillus/Staphylococcus group; Bacillus.  
[1]  
SEQUENCE FROM N.A.  
RA MEDLINE; 94259294.  
RA Hueck C., Kraus A., Hillen W.;  
RT "Sequences of ccpA and two downstream Bacillus megaterium genes with  
RT homology to the motAB operon from Bacillus subtilis.";  
RL Gene 143:147-148(1994).  
CC -!- FUNCTION: MAY BE INVOLVED IN SOME TRANSPORT FUNCTION.  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE MOTB FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; L26052; AAA22297.1; -  
DR PFAM; PF00691; OmpA; 1.  
KW Hypothetical protein; Transsport; Transmembrane.  
FT DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 15 35 POTENTIAL.  
FT DOMAIN 36 218 EXTRACELLULAR (POTENTIAL).  
SQ SEQUENCE 218 AA; 24642 MW; 63622D730AAE247 CRC64;  
  
Query Match 82.5%; Score 47; DB 1; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.99e+01;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
Db 152 AYP5NW 157  
QY 198 TYP5DW 203  
|||||  
-----  
RESULT 3  
ID ECHR\_ERWCH STANDARD; PRT; 250 AA.  
AC Q46967;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE ECHR REGULATORY PROTEIN.  
GN ECHR.  
OS Erwinia chrysanthemi.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Erwinia.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-NCPPB 1066;  
RA Sebatia M., Harrison O., Keil C., Minton N., Salmond G.P.C.;  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: FUNCTIONS AS A POTENTIAL OHLL RESPONSIVE TRANSCRIPTIONAL  
CC REGULATOR.  
CC -!- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL  
CC REGULATORS. AUTOINDUCER REGULATED SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; U45854; AAA66840.1; -  
DR PRINTS; PR00038; HTHLUXR.  
DR PROSITE; PS00622; HTH\_LUXR\_FAMILY; 1.  
DR PFAM; PF00196; Gere; 1.  
KW transcription regulation; DNA-binding; Activator;  
KW Autoinducer induction.

US-09-376-430-2-06.rsp

Thu May 11 06:49:38 2000

DE PROBABLE PROCESSING AND TRANSPORT PROTEIN UL56 (HFLFO PROTEIN).  
GN UL56.  
OS Human cytomegalovirus (strain AD169).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Cytomegalovirus.  
RN [1]  
RX SEQUENCE FROM N.A.  
RX MEDLINE; 90269039.  
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,  
RA Horsnell T., Hutchison C.A. III, Kuzairides T., Martignetti J.A.,  
RA Reddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.:  
RA "Analysis of the protein-coding content of the sequence of human  
cytomegalovirus strain AD169".  
RT Curr. Top. Microbiol. Immunol. 154:125-169(1990).  
RL  
CC -!- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLLOCATION OF THE VIRUS  
CC GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION.  
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL28,  
CC EBV-1 32, PRV ICP18.5, EBV BALF3, AND VZV 30.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X17403; CAA35371.1; -  
CC PIR; S09819; WMBE56.  
CC PFAM; PF01366; PRTP; 1.  
CC Capsid assembly.  
KW SEQUENCE 850 AA; 95868 MW; C32A91906D4FFE7D CRC64;  
SQ  
Query Match 82.5%; Score 47; DB 1; Length 850;  
Best Local Similarity 80.0%; Pred. No. 1.99e+01;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
DB 667 YPSEW 671  
|||:  
QY 199 YPSDW 203  
  
RESULT 6  
ID YP71\_ECOLI STANDARD; PRT; 198 AA.  
AC P21323;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-MAY-1991 (Rel. 18, Last sequence update)  
DT 01-MAY-1991 (Rel. 18, Last annotation update)  
DE HYPOTHETICAL 21.9 KD PROTEIN (ORF1) (RETRON EC67).  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
RN [1]  
RX SEQUENCE FROM N.A.  
RX STRAIN-CL-1;  
RX MEDLINE; 91067724.  
RA Hsu M.-Y., Inouye M., Inouye S.;  
RA "Retron for the 67-base multicopy single-stranded DNA from  
Escherichia coli: a potential transposable element encoding both  
reverse transcriptase and Dam methylase functions."  
RT Proc. Natl. Acad. Sci. U.S.A. 87:9454-9458(1990).  
RL  
CC -!- SIMILARITY: 30% IDENTITY TO CI OF BACTERIOPHAGE 186.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M55249; AAA23390.1; -  
CC PIR; JQ0864; JQ0864.  
DR

KW Transposable element; Hypothetical protein.  
SQ SEQUENCE 198 AA; 21898 MW; 14CED3124E28516B CRC64;  
Query Match 80.7%; Score 46; DB 1; Length 198;  
Best Local Similarity 66.7%; Pred. No. 3.06e+01;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
DB 56 TTPDW 61  
|||:  
QY 198 TTPDW 203  
  
RESULT 7  
ID ST2\_MOUSE STANDARD; PRT; 337 AA.  
AC P14719;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-FEB-2000 (Rel. 39, Last annotation update)  
DE ST2 PROTEIN PRECURSOR (T1 PROTEIN) (LYMPHOCYTE ANTIGEN 84).  
GN ST2 OR STE2 OR LY84.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RX SEQUENCE FROM N.A.  
RX STRAIN-BALB/C;  
RX MEDLINE; 90092495.  
RA Tomlinaga S.;  
RA "A putative protein of a growth specific cDNA from BALB/c-3T3 cells  
is highly similar to the extracellular portion of mouse interleukin 1  
receptor."  
RT FEBS Lett. 258:301-304(1989).  
RL  
RN  
RX SEQUENCE FROM N.A.  
RX STRAIN-C3H/HE; TISSUE-SPLEEN;  
RX MEDLINE; 91355215.  
RA Tomlinaga S.I., Jenkins N.A., Gilbert D.J., Copeland N.G.;  
RA "Molecular cloning of the murine ST2 gene. Characterization and  
chromosomal mapping."  
RT Chromosom. Biophys. Acta 1090:1-8(1991).  
RL  
RN [3]  
RX SEQUENCE FROM N.A.  
RX MEDLINE; 89345536.  
RA Klemenz R., Hoffmann S., Werenskiold A.K.;  
RA "Serum- and oncoprotein-mediated induction of a gene with sequence  
similarity to the gene encoding carcinoembryonic antigen."  
RT Proc. Natl. Acad. Sci. U.S.A. 86:5708-5712(1989).  
RL  
CC -!- FUNCTION: POSSIBLY INVOLVED IN REGULATION OF T-LYMPHOCYTE  
CC ACTION.  
CC -!- DEVELOPMENTAL STAGE: GROWTH-SPECIFIC EXPRESSION, G1-PHASE OF  
CC CELL CYCLE.  
CC -!- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
CC -!- SIMILARITY: STRONG, TO INTERLEUKIN-1 RECEPTORS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Y07519; CAA68812.1; -  
CC EMBL; X60184; CAA42742.1; -  
CC EMBL; M24843; AAA40160.1; -  
CC PIR; S07054; S07054.  
CC MGD; MGI:98427; LY84.  
CC PFAM; PF00047; Ig; 2.  
CC Immunoglobulin domain; Glycoprotein; Signal.  
CC SIGNAL 1 26  
CC CHAIN 27 337  
CC DOMAIN 35 100  
CC  
CC POTENTIAL.  
CC ST2 PROTEIN.  
CC IG-LIKE C2-TYPE DOMAIN.  
DR

FT DOMAIN 132 194 IG-LIKE C2-TYPE DOMAIN.  
 FT DOMAIN 233 315 IG-LIKE C2-TYPE DOMAIN.  
 FT DISULFID 42 93 BY SIMILARITY.  
 FT DISULFID 139 187 BY SIMILARITY.  
 FT DISULFID 240 308 BY SIMILARITY.  
 FT CARBOHYD 60 60 POTENTIAL.  
 FT CARBOHYD 101 101 POTENTIAL.  
 FT CARBOHYD 107 107 POTENTIAL.  
 FT CARBOHYD 146 146 POTENTIAL.  
 FT CARBOHYD 176 176 POTENTIAL.  
 FT CARBOHYD 194 194 POTENTIAL.  
 FT CARBOHYD 225 225 POTENTIAL.  
 FT CARBOHYD 259 259 POTENTIAL.  
 FT CARBOHYD 278 278 POTENTIAL.  
 FT VARIANT 192 192 A -> V (IN STRAIN C3H/HE).  
 SQ SEQUENCE 337 AA; 38502 MW; 7574372722486926 CRC64;

Query Match 80.7%; Score 46; DB 1; Length 337;  
 Best Local Similarity 66.7%; Pred. No. 3.06e+01;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 49 TYPVEW 54  
 ||| :|  
 QY 198 TYPSPDW 203

RESULT 8 STANDARD; PRT; 375 AA.  
 ID TGT\_SHIFL  
 AC Q34117;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE QUEUINE TRNA-RIBOSYLTRANSFERASE (EC 2.4.2.29) (TRNA-GUANINE  
 DE TRANSGLYCOSYLASE) (GUANINE INSERTION ENZYME) (VIRULENCE-ASSOCIATED  
 DE PROTEIN VACC).  
 TGT OR VACC.  
 OS Shigella flexneri.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Shigella.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2A / YSH6000;  
 RX MEDLINE; 94321334.  
 RA Durand J.M., Okada N., Tobe T., Watarai M., Fukuda I., Suzuki T.,  
 RA Nakata N., Komatsu K., Yoshikawa M., Sasakawa C.,  
 RT "vacc, a virulence-associated chromosomal locus of Shigella flexneri,  
 RT is homologous to tgt, a gene encoding trna-guanine transglycosylase  
 RT (tgt) of Escherichia coli K-12.";  
 RL J. Bacteriol. 176:4627-4634(1994).  
 CC -!- FUNCTION: EXCHANGES THE GUANINE RESIDUE WITH 7-AMINOMETHYL-7-  
 CC DEAZAGUANINE IN TRNAS WITH GU(N) ANTICODONS (TRNA-ASP, -ASN, -HIS  
 CC AND -TYR). AFTER THIS EXCHANGE, A CYCLOPENTENDIOL MOIETY IS  
 CC ATTACHED TO THE 7-AMINOMETHYL GROUP OF 7-DEAZAGUANINE, RESULTING  
 CC IN THE HYPERMODIFIED NUCLEOSIDE QUEUOSINE (Q) (7-((4,5-CIS-  
 CC DIHYDROXY-2-CYCLOPENTEN-1-YL)AMINO)METHYL)-7-DEAZAGUANOSINE) (BY  
 CC SIMILARITY).  
 CC -!- FUNCTION: IMPORTANT FOR VIRULENCE.  
 CC -!- CATALYTIC ACTIVITY: TRNA GUANINE + QUEUINE -> TRNA QUEUINE +  
 CC GUANINE.  
 CC -!- COFACTOR: BINDS AND REQUIRES ZINC FOR ACTIVITY. ALSO REQUIRES  
 CC MAGNESIUM (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE QUEUINE TRNA-RIBOSYLTRANSFERASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D26469; BAA05482.1; -  
 DR HSSP; P28720; 1WKD.

KW Queuosine biosynthesis; Transferase; Glycosyltransferase;  
 KW tRNA processing; Zinc; Magnesium; Virulence.  
 FT ACT\_SITE 90 90 BY SIMILARITY.  
 FT ACT\_SITE 265 265 BY SIMILARITY.  
 FT METAL 302 302 ZINC (BY SIMILARITY).  
 FT METAL 304 304 ZINC (BY SIMILARITY).  
 FT METAL 307 307 ZINC (BY SIMILARITY).  
 FT METAL 333 333 ZINC (BY SIMILARITY).  
 SQ SEQUENCE 375 AA; 42607 MW; 52BA93AD0F170720 CRC64;

Query Match 80.7%; Score 46; DB 1; Length 375;  
 Best Local Similarity 80.0%; Pred. No. 3.06e+01;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 148 YPADW 152  
 ||| :|  
 QY 199 YPSDW 203

RESULT 9 STANDARD; PRT; 375 AA.  
 ID TGT\_ECOLI  
 AC P19675; P19676; P78226;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE QUEUINE TRNA-RIBOSYLTRANSFERASE (EC 2.4.2.29) (TRNA-GUANINE  
 DE TRANSGLYCOSYLASE) (GUANINE INSERTION ENZYME).  
 GN TGT.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE; 91177815.  
 RA Reuter K., Slany R., Ullrich F., Kersten H.;  
 RT "Structure and organization of Escherichia coli genes involved in  
 RT biosynthesis of the deazaguanine derivative queuine, a nutrient  
 RT factor for eukaryotes.";  
 RL J. Bacteriol. 173:2256-2264(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE; 97426617.  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1232-1244(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,  
 RA Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,  
 RA Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP PRELIMINARY SEQUENCE OF 168-375 FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE; 91006014.  
 RA Gardel G., Johnson K., Jacq A., Beckwith J.;  
 RT "The sed locus of E. coli codes for two membrane proteins required  
 RT for protein export.";  
 RL EMBO J. 9:3209-3216(1990).  
 RN [5]  
 RP SEQUENCE FROM N.A.; SEQUENCE OF 1-9, AND CHARACTERIZATION.  
 RX MEDLINE; 93287116.  
 RA Garcia G.A., Koch K.A., Chong S.;  
 RT "trna-guanine transglycosylase from Escherichia coli. Overexpression,  
 RT purification and quaternary structure.";  
 RL J. Mol. Biol. 231:489-497(1993).  
 RN [6]  
 RP SEQUENCE OF 281-375 FROM N.A.

FT	MUTAGEN	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773
----	---------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Query Match 80.7%; Score 46; DB 1; Length 436;  
 Best Local Similarity 86.7%; Pred. No. 3.06e+01;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 8 TYPDNW 13  
 ||| :|  
 QY 198 TYPDSW 203

RESULT 11  
 ID EF1H\_XENLA STANDARD; PRT; 437 AA.  
 AC Q91375;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE ELONGATION FACTOR 1-GAMMA TYPE 2 (EF-1-GAMMA) (P47).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;  
 OC Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae;  
 OC Xenopus.  
 RN [1]

SEQUENCE FROM N.A.  
 RP TISSUE-OCYTE;  
 RX MEDLINE; 94155465.  
 RA Morales J., Bassez T., Cormier P., Mulner-Lorillon O., Belle R.,  
 Osborne H.B.;  
 RT "Expression of elongation factor 1 alpha (EF-1 alpha) and 1 beta  
 gamma (EF-1 beta gamma) are uncoupled in early Xenopus embryos.";  
 RL Dev. Genet. 14:440-448(1993).  
 CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN ANCHORING THE COMPLEX TO OTHER  
 CELLULAR COMPONENTS.  
 CC -1- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,  
 CC DELTA, AND GAMMA.  
 CC -1- SIMILARITY: TO OTHER SPECIES' GAMMA SUBUNIT AS WELL AS TO THE  
 CC N-TERMINAL OF SOME EUKARYOTIC AMINO ACID-TRNA SYNTHETASES.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; S69726; AAB29958.1;  
 DR PROSITE; PS50040; EF1G; 1.  
 DR PFAM; PF00043; GST; 1.  
 DR PFAM; PF00647; EF1G\_domain; 1.  
 KW Elongation factor; Protein biosynthesis.  
 SQ SEQUENCE 437 AA; 50248 MW; 9DB03AD4FC87E557 CRC64;

Query Match 80.7%; Score 46; DB 1; Length 437;  
 Best Local Similarity 66.7%; Pred. No. 3.06e+01;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 8 TYPDNW 13  
 ||| :|  
 QY 198 TYPDSW 203

RESULT 12  
 ID ACHD\_HUMAN STANDARD; PRT; 517 AA.  
 AC Q07001;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE ACETYLCHOLINE RECEPTOR PROTEIN, DELTA CHAIN PRECURSOR.  
 GN CHRD OR ACHRD.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89177471.  
 RA Luther M.A., Schoepfer R., Whiting P., Casey B., Blatt Y.,  
 RA Montal M.S., Montal M., Lindstrom J.;  
 RT "A muscle acetylcholine receptor is expressed in the human cerebellar  
 medulloblastoma cell line TE671.";  
 RL J. Neurosci. 9:1082-1096(1989).  
 CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN  
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
 CC MEMBRANE.  
 CC -1- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,  
 CC DELTA, AND GAMMA (IN IMMATURE MUSCLE) OR EPSILON (IN MATURE  
 CC MUSCLE) CHAINS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X55019; CAA38759.1;  
 DR PIR; A60916; A60916.  
 DR MIM; 100720;  
 DR PRINTS; PR00252; NRIONCHANNEL.  
 DR PRINTS; PR00254; NICOTINIC.  
 DR PROSITE; PS00236; NEUROTR\_ION\_CHANNEL; 1.  
 DR PFAM; PF00065; neur\_chan; 1.  
 KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
 KW Transmembrane; Phosphorylation.  
 FT SIGNAL 1 21 BY SIMILARITY.  
 FT CHAIN 22 517 ACETYLCHOLINE RECEPTOR PROTEIN, DELTA  
 FT DOMAIN 22 245  
 FT TRANSMEM 246 270 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 278 299 POTENTIAL.  
 FT TRANSMEM 312 333 POTENTIAL.  
 FT DOMAIN 334 471 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 472 490 POTENTIAL.  
 FT DISULFID 151 165 BY SIMILARITY.  
 FT CARBOHYD 97 97 POTENTIAL.  
 FT CARBOHYD 164 164 POTENTIAL.  
 FT MOD\_RES 390 390 PHOSPHORYLATION (BY TYR-KINASES)  
 FT (BY SIMILARITY).  
 SQ SEQUENCE 517 AA; 58895 MW; 195CEF69358758BD CRC64;

Query Match 80.7%; Score 46; DB 1; Length 517;  
 Best Local Similarity 66.7%; Pred. No. 3.06e+01;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 192 TYPVIEW 197  
 ||| :|  
 QY 198 TYPDSW 203

RESULT 13  
 ID Y414\_MYCGF STANDARD; PRT; 1036 AA.  
 AC P47653; P47654; Q49457;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HYPOTHETICAL PROTEIN MG414.  
 GN MG414.  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;

sequence, amino-acid sequence of proteolytic peptides of the enzyme and comparison of the structure to those of other known aminoacyl-tRNA synthetases."

Biol. Chem. Hoppe-Seyler 368:971-979(1987).

[3]

SEQUENCE FROM N.A.

STRAIN-S288C

MEDLINE; 96076635.

RA Obermaier B., Gassenhuber J., Piravandi E., Domdey H.;

RT "Sequence analysis of a 78.6 kb segment of the left end of

RT Saccharomyces cerevisiae chromosome II."

RL Yeast 11:1103-1112(1995).

[4]

RN SEQUENCE FROM N.A.

RC STRAIN-S288C

RA Contreras R., Fiers W., Logghe M., Molemans F.;

RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: ATP + L-ISOLEUCINE + TRNA(ILE) = AMP +

CC -1- PHOSPHORYLATION: L-ISOLEUCYL-TRNA(ILE).

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC

CC -1- SIMILARITY: BELONGS TO CLASS-1 AMINOACYL-TRNA SYNTHETASE FAMILY

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; X07886; CAA30733.1;

DR EMBL; M19992; AAA34712.1;

DR EMBL; X79489; CAA56034.1;

DR EMBL; Z35838; CAA84898.1;

DR PIR; S14459; SYBY14.

DR PIR; S05761; SYBY11.

DR SGD; L0000856; ILS1.

DR PROSITE; PRO0984; TRNASYNTHILE.

DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_1; 1.

DR PFAM; PF00133; trna-synt.1; 1.

KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.

FT SIMILAR 47 57

FT BINDING 602 606

FT BINDING 605 605

FT CONFLICT 551 551

FT CONFLICT 588 588

FT CONFLICT 693 693

FT CONFLICT 706 706

FT CONFLICT 747 747

FT CONFLICT 1072 1072

FT CONFLICT 122983 122983

FT CONFLICT 75E07F3B0A4DBBF CRC64;

FT SEQUENCE 1072 AA; 122983 MW; 75E07F3B0A4DBBF CRC64;

Query Match 80.7%; Score 46; DB 1; Length 1072;

Best Local Similarity 83.3%; Pred. No. 3.06e+01;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 125 TYASDW 130

QY 198 TYPDSW 203

RESULT 15

ID VIT\_BOMMO STANDARD; PRT; 1782 AA.

AC 027309;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE VITELLOGENIN PRECURSOR [CONTAINS: VITELLIN LIGHT CHAIN (VL); VITELLIN HEAVY CHAIN (VH)].

DE

GN

OS Bombyx mori (Silk moth).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Prorygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;

OC Bombycoidea; Bombycidae; Bombyx.

MEDLINE; 96026346.

RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,

RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,

RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.D.,

RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,

RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,

RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;

RT "The minimal gene complement of Mycoplasma genitalium."

RL Science 270:397-403(1995).

[2]

RN SEQUENCE OF 52-146 AND 733-833 FROM N.A.

RC STRAIN-A1000 32530 / G-37;

RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;

RT "A survey of the Mycoplasma genitalium genome by using random

RT sequencing."

RL J. Bacteriol. 175:7918-7930(1993).

CC -1- SIMILARITY: BELONGS TO THE MG414 / MG415 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; U39723; AAC71641.1;

DR EMBL; U01695; AAB01008.1;

DR EMBL; U01804; AAD12330.1;

DR TIGR; MG414;

KW Hypothetical protein; Transmembrane.

FT TRANSMEM 4 24

FT TRANSMEM 1004 1024

FT CONFLICT 733 736

FT CONFLICT 1036 1036

FT SEQUENCE 1036 AA; 123179 MW; 12A21F00F686A141 CRC64;

Query Match 80.7%; Score 46; DB 1; Length 1036;

Best Local Similarity 57.1%; Pred. No. 3.06e+01;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 79 TYSNEWS 85

QY 198 TYPDSWS 204

RESULT 14

ID SYIC\_YEAST STANDARD; PRT; 1072 AA.

AC P09436;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 01-FEB-1996 (Rel. 33, Last annotation update)

DE ISOLEUCYL-TRNA SYNTHETASE, CYTOPLASMIC (EC 6.1.1.5) (ISOLEUCINE--TRNA LIGASE) (ILERS).

DE

GN ILS1 OR YBL076C OR YBL0734.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;

OC Saccharomycetaceae; Saccharomyces.

RN

RP SEQUENCE FROM N.A.

RC STRAIN-A364N;

RA Martindale D.W., Gu Z.M.;

RT "Isolation and complete sequence of the yeast isoleucyl-tRNA

RT synthetase gene (ILS1)."

RL Curr. Genet. 15:99-106(1989).

[2]

RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC STRAIN-S288C;

RC MEDLINE; 88024443.

RA Englisch U., Englisch E., Markmeyer P., Schischkoff J.,

RA Sternbach H., Kratzin H., Gramer F.;

RT "Structure of the yeast isoleucyl-tRNA synthetase gene (ILS1). DNA-





\*\*\*\*\*  
 WPSRCH\_PP  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 11:48:18 2000; MasPar time 223.70 Seconds  
 Tabular output not generated. 2.170 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (198-204) from US09376430A.pep (6 of 25)  
 Perfect Score: 57  
 Sequence: 1 TYPDWS 7

Scoring table: PAM 150  
 Gap 11

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: sptrembl12  
 1:sp\_archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
 5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
 9:sp\_phase 10:sp\_plant 11:sp\_rodent 12:sp\_unclassified  
 13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 22.932; Variance 28.140; scale 0.815

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	55	96.5	486	11	CYTOSOLIC PHOSPHOPRO	7.87e-01
2	52	91.2	335	4	DJ437M21.3 (NOVEL SRC	3.48e+00
3	52	91.2	448	13	P52.	3.48e+00
4	49	86.0	112	1	ORF H0011.	1.46e+01
5	49	86.0	242	2	TRANSRIPTONAL ACTIVA	1.46e+01
6	49	86.0	293	2	GLYCINEA AVRD.	1.46e+01
7	49	86.0	311	2	PV. LACHRYMANS (AVRD)	1.46e+01
8	49	86.0	311	2	AVIRULENCE D PROTEIN.	1.46e+01
9	49	86.0	311	2	PV. PHASEOLICOLA 3121	1.46e+01
10	49	86.0	311	2	AVIRULENCE D PROTEIN.	1.46e+01
11	49	86.0	311	2	PHASEOLICOLA AVRD.	1.46e+01
12	49	86.0	311	2	GLYCINEA AVRD.	1.46e+01
13	49	86.0	311	2	GLYCINEA AVRD.	1.46e+01
14	49	86.0	311	2	PV. LACHRYMANS (AVRD)	1.46e+01
15	49	86.0	506	2	PUTATIVE TRNA SYNTHETA	1.46e+01
16	48	84.2	166	1	HYPOTHETICAL 18.3 KD P	2.32e+01
17	48	84.2	230	10	F12L6.13 PROTEIN.	2.32e+01
18	48	84.2	247	2	RESPONSE REGULATOR YSP	2.32e+01
19	48	84.2	419	2	HYPOTHETICAL 47.0 KD P	2.32e+01
20	48	84.2	476	2	NOEA.	2.32e+01

21	47	82.5	229	2	087971	RESPONSE REGULATOR YUK	3.67e+01
22	47	82.5	287	1	Q9YEP5	287AA LONG HYPOTHETICA	3.67e+01
23	47	82.5	439	3	Q13295	ALPHA-GALACTOSIDASE.	3.67e+01
24	47	82.5	441	11	Q61644	H74 PROTEIN.	3.67e+01
25	47	82.5	470	14	Q9WJ27	HELFO (FRAGMENT).	3.67e+01
26	47	82.5	621	12	Q59301	ENDO-BETA-1,4-XYLANASE	3.67e+01
27	47	82.5	850	14	Q42234	UL56 PROTEIN PRECURSOR	3.67e+01
28	47	82.5	1568	3	Q74415	HYPOTHETICAL 179.2 KD	3.67e+01
29	46	80.7	102	1	Q59294	102AA LONG HYPOTHETICA	3.67e+01
30	46	80.7	176	2	Q54412	PARTICULATE METHANE MO	5.77e+01
31	46	80.7	201	2	Q92IH7	CP52P.	5.77e+01
32	46	80.7	220	2	Q68537	B5CK.	5.77e+01
33	46	80.7	258	2	Q9X4J0	PROZL.	5.77e+01
34	46	80.7	266	2	Q75895	ECHE PROTEIN.	5.77e+01
35	46	80.7	410	5	Q02341	ZK896.1 PROTEIN.	5.77e+01
36	46	80.7	422	4	Q16542	INTERLEUKIN-11 RECEPTO	5.77e+01
37	46	80.7	435	10	Q40452	A MEMBRANE-ASSOCIATED	5.77e+01
38	46	80.7	524	5	Q01321	F26H9.1 PROTEIN.	5.77e+01
39	46	80.7	567	11	Q05208	ST2L PROTEIN PRECURSOR	5.77e+01
40	46	80.7	592	2	Q05338	HYPOTHETICAL 59.8 KD P	5.77e+01
41	46	80.7	798	2	Q08329	GLYCOGEN PHOSPHORYLASE	5.77e+01
42	46	80.7	978	3	Q94258	PUTATIVE PRE-TENA NUCL	5.77e+01
43	46	80.7	983	3	Q94545	POSSIBLE RANBP7-IMPORT	5.77e+01
44	46	80.7	1067	14	Q69904	FROM AFRICAN GREEN MON	5.77e+01
45	46	80.7	2548	14	Q9WRN0	TEGUMENT PROTEIN.	5.77e+01

## ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	486 AA.
ID	Q9WVE8			
AC	Q9WVE8			
DT	01-NOV-1999	(TREMELrel. 12, Created)		
DT	01-NOV-1999	(TREMELrel. 12, Last sequence update)		
DT	01-NOV-1999	(TREMELrel. 12, Last annotation update)		
DE	CYTOSOLIC PHOSPHOPROTEIN PACSIN2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL6 X DBA;			
RA	RITTER B., MODREGGER J., PAULSSON M., PLOMANN M.;			
RT	"PACSIN2, a novel member of the PACSIN family of cytoplasmic			
RT	phosphoproteins."			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AFL28535; AAD41780.1;			
SQ	SEQUENCE 486 AA; 55833 MW; F57E5FAA CRC32;			

Query Match 96.5%; Score 55; DB 11; Length 486;  
 Best Local Similarity 85.7%; Pred. No. 7.87e-01;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db	393	TYPDWS	399	
QY	198	TYPDWS	204	
RESULT	2	PRELIMINARY:	PRT:	335 AA.
ID	Q9Y4V2			
AC	Q9Y4V2			
DT	01-NOV-1999	(TREMELrel. 12, Created)		
DT	01-NOV-1999	(TREMELrel. 12, Last sequence update)		
DT	01-NOV-1999	(TREMELrel. 12, Last annotation update)		
DE	DJ437M21.3 (NOVEL SRC HOMOLOG DOMAIN 3 CONTAINING PROTEIN SIMILAR TO			
DE	MOUSE H74 AND CHICKEN PAP52) (FRAGMENT).			
GN	DJ437M21.3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Primates; Catarrhini; Hominidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	LAIRD G.;			

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL049758; CAB51395.1; -  
 FT NON-TER  
 SQ SEQUENCE 335 AA; 37945 MW; 2A09BAD3 CRC32;

Query Match 91.2%; Score 52; DB 4; Length 335;  
 Best Local Similarity 71.4%; Pred. No. 3.48e+00;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Db 242 SYPTDWS 248  
 QY 198 TYPDWS 204

RESULT 3  
 ID O13154 PRELIMINARY; PRT; 448 AA.  
 AC O13154;  
 DT 01-JUL-1997 (T-EMBLrel. 04, Created)  
 DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)  
 DT 01-NOV-1999 (T-EMBLrel. 12, Last annotation update)  
 DE P52  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;  
 OC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE; 97435298.  
 RA MERILAINEN J., LEHTO V.P., WASENIUS V.M.;  
 RT "FAP52, a novel, SH3 domain-containing focal adhesion protein.";  
 RL J. Biol. Chem. 272:23278-23284(1997).  
 DR EMBL; 250798; CAA90678.1;  
 DR HSP; 060631; 4GBQ.  
 DR PFAM; PF00611; FGH; 1.  
 DR PFAM; PF00018; SH3; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 SQ SEQUENCE 448 AA; 51971 MW; 5B3C6CBD CRC32;

Query Match 91.2%; Score 52; DB 13; Length 448;  
 Best Local Similarity 71.4%; Pred. No. 3.48e+00;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Db 355 SYPTDWS 361  
 QY 198 TYPDWS 204

RESULT 4  
 ID O51956 PRELIMINARY; PRT; 112 AA.  
 AC O51956;  
 DT 01-JUN-1998 (T-EMBLrel. 06, Created)  
 DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)  
 DT 01-JUN-1998 (T-EMBLrel. 06, Last annotation update)  
 DE ORF H0011.  
 OS Halobacterium sp.  
 OG Plasmid pNRC100.  
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;  
 OC Halobacterium.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-NRC-1;  
 RX MEDLINE; 88201675.  
 RA DASSARMA S., DAMERVAL T., JONES J.G., TANDEAU DE MARSAC N.;  
 RT "A plasmid-encoded gas vesicle protein gene in a halophilic archaeobacterium";  
 RL Mol. Microbiol. 1:365-370(1987).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-NRC-1;  
 RX MEDLINE; 90016863.  
 RA JONES J.G., HACKETT N.R., HALLADAY J.T., SCOTHORN D.J., YANG C.F.,  
 RA NG W.L., DASSARMA S.;  
 RT "Analysis of insertion mutants reveals two new genes in the pNRC100

RT gas vesicle gene cluster of Halobacterium halobium.";  
 RL Nucleic Acids Res. 17:7785-7793(1989).  
 RN [3]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-NRC-1;  
 RX MEDLINE; 91267967.  
 RA NG W.L., KOTHAKOTA S., DASSARMA S.;  
 RT "Structure of the gas vesicle plasmid in Halobacterium halobium  
 RT inversion isomers, inverted repeats, and insertion sequences.";  
 RL J. Bacteriol. 173:3933-3933(1991).  
 RN [4]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-NRC-1;  
 RX MEDLINE; 91323716.  
 RA JONES J.G., YOUNG D.C., DASSARMA S.;  
 RT "Structure and organization of the gas vesicle gene cluster on the  
 RT Halobacterium halobium plasmid pNRC100.";  
 RL Gene 102:117-122(1991).  
 RN [5]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-NRC-1;  
 RX MEDLINE; 93012964.  
 RA HALLADAY J.T., NG W.L., DASSARMA S.;  
 RT "Genetic transformation of a halophilic archaeobacterium with a gas  
 RT vesicle gene cluster restores its ability to float.";  
 RL Gene 119:131-136(1992).  
 RN [6]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-NRC-1;  
 RX MEDLINE; 93139036.  
 RA HALLADAY J.T., JONES J.G., LIN F., MACDONALD A.B., DASSARMA S.;  
 RT "The rightward gas vesicle operon in Halobacterium plasmid pNRC100:  
 RT identification of the gvpA and gvpC gene products by use of antibody  
 RT probes and genetic analysis of the region downstream of gvpC.";  
 RL J. Bacteriol. 175:684-692(1993).  
 RN [7]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-NRC-1;  
 RX MEDLINE; 93327890.  
 RA DASSARMA S.;  
 RT "Identification and analysis of the gas vesicle gene cluster on an  
 RT unstable plasmid of Halobacterium halobium.";  
 RL Experientia 49:482-486(1993).  
 RN [8]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-NRC-1;  
 RX MEDLINE; 93328662.  
 RA NG W.L., DASSARMA S.;  
 RT "Minimal replication origin of the 200-kilobase Halobacterium plasmid  
 RT pNRC100.";  
 RL J. Bacteriol. 175:4584-4596(1993).  
 RN [9]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-NRC-1;  
 RX MEDLINE; 95095934.  
 RA DASSARMA S., ARORA P., LIN F., MOLINARI E., YIN L.R.;  
 RT "Wild-type gas vesicle formation requires at least ten genes in the  
 RT gvp gene cluster of Halobacterium halobium plasmid pNRC100.";  
 RL J. Bacteriol. 176:7646-7652(1994).  
 RN [10]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-NRC-1;  
 RX MEDLINE; 16:560-568(1994).  
 RA NG W.L., ARORA P., DASSARMA S.;  
 RT Syst. Appl. Microbiol. 16:560-568(1994).  
 RN [11]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-NRC-1;  
 RX MEDLINE; 90016863.  
 RA JONES J.G., HACKETT N.R., HALLADAY J.T., SCOTHORN D.J., YANG C.F.,  
 RA NG W.L., DASSARMA S.;  
 RT "Analysis of insertion mutants reveals two new genes in the pNRC100



QY 198 TYPDWS 204

RESULT 9  
ID Q52530 PRELIMINARY; PRT; 311 AA.  
AC Q52530;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE PV. PHASEOLICOLA 3121 (AVRD) (FRAGMENT).  
GN AVRD.  
OS Pseudomonas syringae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;  
OC Pseudomonas.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-3121;  
RX MEDLINE; 94220735.  
RA YUCEL I., BOYD C., DEBNAM Q., KEEN N.T.;  
RT "Two different classes of avrd alleles occur in pathovars of  
RT Pseudomonas syringae."  
RL Mol. Plant Microbe Interact. 7:131-139(1994).  
DR EMBL; L11336; AAA20579.1;  
FT NON\_TER 311 311  
SQ SEQUENCE 311 AA; 34778 MW; CEE7B4E0 CRC32;

Query Match 86.0%; Score 49; DB 2; Length 311;  
Best Local Similarity 57.1%; Pred. No. 1.46e+01;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 49 SYPANWS 55  
QY 198 TYPDWS 204

RESULT 10  
ID O86072 PRELIMINARY; PRT; 311 AA.  
AC O86072;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE AVIRULENCE D PROTEIN.  
GN AVRD.  
OS Pseudomonas syringae pv. apii.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;  
OC Pseudomonas.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA KEEN N.T., BOYD C.M.;  
RL Submitted (AUG-1998), to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF083919; AAC33120.1;  
SQ SEQUENCE 311 AA; 34768 MW; 8EBB0511 CRC32;

Query Match 86.0%; Score 49; DB 2; Length 311;  
Best Local Similarity 57.1%; Pred. No. 1.46e+01;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 49 SYPANWS 55  
QY 198 TYPDWS 204

RESULT 11  
ID O08244 PRELIMINARY; PRT; 311 AA.  
AC O08244;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
DE PHASEOLICOLA AVRD.  
GN AVRD.  
OS Pseudomonas syringae (pv. phaseolicola).  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;  
OC Pseudomonas.

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RACE 2;  
RA KEITH L.M., BOYD C., KEEN N.T., PARTRIDGE J.E.;  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U87228; AAB53626.1;  
SQ SEQUENCE 311 AA; 34715 MW; D7D7341E CRC32;

Query Match 86.0%; Score 49; DB 2; Length 311;  
Best Local Similarity 57.1%; Pred. No. 1.46e+01;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 49 SYPANWS 55  
QY 198 TYPDWS 204

RESULT 12  
ID O08243 PRELIMINARY; PRT; 311 AA.  
AC O08243;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
DE GLYCINEA AVRD.  
GN AVRD.  
OS Pseudomonas syringae (pv. glycinea).  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;  
OC Pseudomonas.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RACE 3, RACE 6;  
RA KEITH L.M., BOYD C., KEEN N.T., PARTRIDGE J.E.;  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U87227; AAB53625.1;  
SQ SEQUENCE 311 AA; 34523 MW; 6A568357 CRC32;

Query Match 86.0%; Score 49; DB 2; Length 311;  
Best Local Similarity 57.1%; Pred. No. 1.46e+01;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 49 SYPANWS 55  
QY 198 TYPDWS 204

RESULT 13  
ID O08242 PRELIMINARY; PRT; 311 AA.  
AC O08242;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
DE GLYCINEA AVRD.  
GN AVRD.  
OS Pseudomonas syringae (pv. glycinea).  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;  
OC Pseudomonas.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RACE 1;  
RA KEITH L.M., BOYD C., KEEN N.T., PARTRIDGE J.E.;  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U87226; AAB53624.1;  
SQ SEQUENCE 311 AA; 34616 MW; 362F7E39 CRC32;

Query Match 86.0%; Score 49; DB 2; Length 311;  
Best Local Similarity 57.1%; Pred. No. 1.46e+01;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 49 SYPANWS 55  
QY 198 TYPDWS 204

Search completed: Wed May 10 11:52:13 2000  
Job time : 235 secs.

```
RESULT 14
ID Q52426 PRELIMINARY; PRT; 311 AA.
AC Q52426;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE PV. LACHRYMANS (AVRD) (FRAGMENT).
OS AVRD.
GN Pseudomonas syringae.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
OC Pseudomonas.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94220735.
RA YUCEL I.; BOYD C.; DEBNAM Q.; KEEN N.T.;
RT "Two different classes of avrD alleles occur in pathovars of
RT Pseudomonas syringae.";
RL Mol. Plant Microbe Interact. 7:131-139(1994).
DR EMBL; L11334; AAA20577.1; -.
FT NON_TER 311
SQ SEQUENCE 311 AA; 34465 MW; 152E6203 CRC32;

Query Match 86.0%; Score 49; DB 2; Length 311;
Best Local Similarity 57.1%; Pred. NO. 1.46e+01;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 49 SYPANWS 55
QY 198 TYPDWS 204
:|:|:|:
```

```
RESULT 15
ID Q92BH7 PRELIMINARY; PRT; 506 AA.
AC Q92BH7;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE PUTATIVE TRNA SYNTHETASE.
GN SC9B5.03.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA SEEGER K.J.; HARRIS D.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA BENTLEY S.D.; PARKHILL J.; BARRELL B.G.; RAJANDREAM M.A.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE; 97000351.
RA REDENBACH M.; KIESER H.M.; DENAPAITE D.; EICHNER A.; CULLUM J.;
RA KINASHI H.; HOPWOOD D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL035206; CAA22745.1; -.
KW Aminoacyl-tRNA synthetase.
SQ SEQUENCE 506 AA; 55170 MW; C8B52255 CRC32;
```

```
Query Match 86.0%; Score 49; DB 2; Length 506;
Best Local Similarity 57.1%; Pred. NO. 1.46e+01;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 236 SYPADWG 242
QY 198 TYPDWS 204
:|:|:|:
```



\*\*\*\*\*

W E R L E H

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 12:01:15 2000; MasPar time 3.74 Seconds  
Tabular output not generated. 221.479 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (226-260) from US09376430A.pep (7 of 25)  
Perfect Score: 275  
Sequence: 1 KPRLSKFILISSAILLMVSLLLSLWKLVRVKKF 35

Scoring table: PAM 150  
Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq35  
1:geneseqp

Statistics: Mean 25.049; Variance 138.308; scale 0.181

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	104	37.8	1183	1 Y07728	Armenian hamster alpha	4.00e+00
2	100	36.4	593	1 W34674	Human mammary carcinoma	7.24e+00
3	100	36.4	876	1 W34575	Human mammary carcinoma	7.24e+00
4	100	36.4	824	1 W34573	Human mammary carcinoma	7.24e+00
5	100	36.4	914	1 R71100	Protein-tyrosine-kinase	7.24e+00
6	100	36.4	919	1 W34572	Human mammary carcinoma	7.24e+00
7	100	36.4	919	1 R75502	Human mammary carcinoma	7.24e+00
8	100	36.4	919	1 R75504	Human mammary carcinoma	7.24e+00
9	99	36.0	1073	1 R28821	Alpha 6A integrin subu	8.39e+00
10	99	36.0	1073	1 R28822	Alpha 6B integrin subu	8.39e+00
11	99	36.0	1091	1 R28822	Alpha 6B integrin subu	8.39e+00
12	96	34.9	505	1 W03761	Mullerian inhibiting s	1.30e+01
13	96	34.9	505	1 R41922	MISR3.	1.30e+01
14	96	34.9	788	1 R26322	B6H.	1.30e+01
15	96	34.9	799	1 W02194	Human integrin beta su	1.30e+01
16	95	34.5	141	1 R28823	Alpha 6B integrin subu	1.51e+01
17	95	34.5	149	1 R28824	Alpha 6A integrin subu	1.51e+01
18	94	34.2	1285	1 W72972	Drosophila melanogaste	1.75e+01
19	94	34.2	1299	1 R86304	Drosophila patched pro	1.75e+01
20	91	33.1	422	1 W88360	Human lymphocyte activ	2.70e+01
21	91	33.1	498	1 R87089	Human immunoglobulin s	2.70e+01
22	91	33.1	498	1 R13270	Lymphocyte Activation	2.70e+01
23	91	33.1	503	1 W27507	Human activin receptor	2.70e+01

24	91	33.1	503	1 R55366	Human Activin receptor	2.70e+01
25	91	33.1	503	1 R85210	Human ALK-1.	2.70e+01
26	91	33.1	503	1 R94602	TAR-3 polypeptide.	2.70e+01
27	90	32.7	319	1 W35855	Human CD43 for use in	3.12e+01
28	89	32.4	130	1 W12723	PHIF truncated gene pr	3.61e+01
29	89	32.4	202	1 W12721	PHIF gene product.	3.61e+01
30	88	32.0	774	1 R22120	Sequence encoded by a	4.17e+01
31	88	32.0	800	1 R07713	Human low density lipo	4.17e+01
32	88	32.0	860	1 R47860	Human LDL receptor.	4.17e+01
33	88	32.0	860	1 R47157	Sequence of human low	4.17e+01
34	88	32.0	924	1 R78234	Chicken p95/human LDL	4.17e+01
35	87	31.6	774	1 R93986	Cephalosporin C acylas	4.81e+01
36	87	31.6	875	1 R73053	Peptidyl C-terminal al	4.81e+01
37	87	31.6	875	1 P94854	C-terminal prepro-C-te	4.81e+01
38	86	31.3	28	1 P91339	Amino acid sequence of	5.55e+01
39	86	31.3	105	1 P50297	Human epidermal growth	5.55e+01
40	86	31.3	383	1 W01564	p(rat-edg), G-protein	5.55e+01
41	86	31.3	383	1 W87791	Rat-edg, G-protein cou	5.55e+01
42	86	31.3	448	1 W53463	Human gp49 Hm18 polype	5.55e+01
43	86	31.3	448	1 W82551	Human LIR-pbm2 protein	5.55e+01
44	86	31.3	472	1 W69234	FCR-IV protein sequenc	5.55e+01
45	86	31.3	761	1 W99790	Rat VRRP-1 (VR2) caps	5.55e+01

## ALIGNMENTS

RESULT 1  
ID Y07728 standard; protein; 1183 AA.

AC Y07728;  
DT 01-JUL-1999 (first entry)  
DE Armenian hamster alpha-1 integrin subunit protein.  
KW VEGF; tumour angiogenesis inhibition; vascular endothelial growth factor;  
KW integrin cell surface receptor; capillary; blood vessel; hamster;  
KW alpha-1 subunit; alpha-2 subunit.  
OS Cricetus migratorius.  
PN W09916465-A1.  
PD 08-APR-1999.  
PR 30-SEP-1997; U17485.  
PR 30-SEP-1997; WO-017485.  
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
PA (CLAF/) CLAFFEY K P.  
PA (DETM/) DETMAR M.  
PA (SENG/) SENG D R.  
PI Claffey KP, Detmar M, Senger DR;  
DR WPI: 99-254930/21.  
PT Inhibition of tumor angiogenesis through interaction of vascular  
PT endothelial growth factor and integrin cell surface receptors  
PS Disclosure: Fig 2A-C: 64pp; English  
CC This invention describes a novel method for the inhibition of tumour  
CC angiogenesis mediated by vascular endothelial growth factor (VEGF) and  
CC integrin cell surface receptors expressed in vasculature of living  
CC subjects. The method inhibits new capillary and new blood vessel  
CC formation both within a tumour mass itself as well as in the immediately  
CC adjacent blood vasculature surrounding the perimeter of the tumour mass.  
CC Interaction and dependence upon VEGF to induce specific integrin  
CC heterodimers in tumour angiogenesis provides a novel method for  
CC inhibiting tumour angiogenesis, and unlike prior art relies on the  
CC specific inter-relationship of VEGF and integrins, rather than  
CC concentrating solely on one specific class of protein.  
SQ Sequence 1183 AA;

Query Match 37.8%; Score 104; DB 1; Length 1183;  
Best Local Similarity 54.5%; Pred. No. 4.00e+00;  
Matches 12; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

Db 1149 ILISAFAGLLMLLLIILALWKI 1170  
QY 233 ILISSUAILLMVSLLLSLWKL 254

RESULT 2  
ID W34674 standard; Protein; 563 AA.  
AC W34674;

17-FEB-1998 (first entry)  
 Human mammary carcinoma kinase 10 (MCK-10) splice variant 2.  
 Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;  
 proliferative disease; cancer; insulin receptor family;  
 tyrosine kinase neurotrophin receptor; MCK-10 activity;  
 neurological disorder; aberrant expression.  
 OS Homo sapiens.  
 FH Key  
 FT Location/Qualifiers  
 FT 1..18  
 FT /label= signal\_sequence  
 FT 19..919  
 FT /note= "mature\_protein"  
 FT 31..185  
 FT /label= discoidin\_I\_like\_domain  
 FT 304..307  
 FT /label= endopeptidase\_furin  
 FT /note= "putative precursor cleavage site"  
 FT 48..439  
 FT /label= transmembrane\_region  
 FT 617..627  
 FT /label= ATP\_binding\_motif  
 FT 797..798  
 FT /label= autophosphorylation\_sites  
 FT /note= "putative"  
 FT 793  
 FT /label= autophosphorylation\_site  
 FT /note= "putative"  
 FT 839..842  
 FT /label= binding\_motif\_for\_P13\_kinase  
 FT /note= "binding motif for phosphatidylinositol 3' kinase"  
 FT 827..827  
 FT /label= potential\_substrate\_binding\_site  
 FT 506..509  
 FT /label= putative\_receptor\_binding\_site\_for\_SHC  
 FT /note= "SHC is an oncogenic SH2 domain containing molecule"  
 FT 510..513  
 FT /label= GTPase\_activity\_protein\_binding\_site  
 FT /note= "putative"  
 FT 26..42  
 FT /note= "antibody recognition sequence NTalpha"  
 FT 309..321  
 FT /note= "antibody recognition sequence NTBeta"  
 FT 897..913  
 FT /note= "antibody recognition sequence CTbeta"  
 US5677144-A.  
 14-OCT-1997.  
 08-NOV-1994; 336343.  
 16-NOV-1993; US-153397.  
 (ALVE/) ALVES F H E.  
 (ULLR/) ULLRICH A.  
 PI Alves FHE, Ullrich A;  
 DR WPI; 97-511869/47.  
 Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding  
 for it, useful for cancer diagnosis  
 PS Disclosure; Page -; 70pp; English.  
 The present sequence represents a splice variant of a mammary  
 carcinoma kinase (MCK-10). This kinase belongs to a novel family  
 of receptor tyrosine kinases, and expression is associated with  
 proliferative diseases such as cancer. The MCK-10 receptor tyrosine  
 kinase has extensive sequence similarity to the insulin receptor family.  
 The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide  
 primer pools, using a template cDNA synthesised by reverse transcription  
 of poly-A RNA from the human mammary carcinoma cell line MCF7. The  
 amplified PCR product was used to screen human foetal brain and placental  
 libraries, from which the present splice variant was isolated. This  
 splice variant does not possess amino acids 666-671 of MCK-10 (W34672).  
 The sequence represented by amino acids 585-595 may be important, as  
 deletion of this motif in the activin receptor serine/threonine kinase  
 results in reduced ligand binding affinity. MCK-10 is expressed in brain  
 tissue, and the protein shares homology with the tyrosine kinase  
 neurotrophin receptor. Modulation of MCK-10 activity therefore may be used

CC for treatment of neurological disorders. MCK-10 is also expressed in a  
 CC variety of cancer cell lines and tumour tissue. The nucleotide sequence  
 CC of MCK-10, or parts of it, can be used for diagnostic purposes to detect  
 CC aberrant expression of MCK-10 genes. Inhibitors of MCK-10 (or splice  
 CC variants) receptor activity may have therapeutic value in the treatment  
 CC of diseases such as cancer.  
 CC note: the present sequence does not appear in the specification, but was  
 CC created using information provided.  
 SQ -sequence 563 AA;  
 Query Match 36.4%; Score 100; DB 1; Length 563;  
 Best Local Similarity 51.9%; Pred.No. 7.24e+00;  
 Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;  
 Db 68 ILTGCLVAIIILLIIALLMLRLHWR 94  
 QY 233 ILTSSL-AILLMVSLLL-LSLWKL-WR 256  
 |||:| ||::|::| ||:| ||  
 RESULT 3  
 ID W34675 standard; Protein; 876 AA.  
 AC W34675;  
 DT 17-FEB-1998 (first entry)  
 DE Human mammary carcinoma kinase 10 (MCK-10) splice variant 3.  
 KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;  
 KW proliferative disease; cancer; insulin receptor family;  
 KW tyrosine kinase neurotrophin receptor; MCK-10 activity;  
 KW neurological disorder; aberrant expression.  
 OS Homo sapiens.  
 FH Key  
 FT Location/Qualifiers  
 FT 1..18  
 FT /label= signal\_sequence  
 FT 19..876  
 FT /note= "mature\_protein"  
 FT 31..185  
 FT /label= discoidin\_I\_like\_domain  
 FT 304..307  
 FT /label= endopeptidase\_furin  
 FT /note= "putative precursor cleavage site"  
 FT 48..439  
 FT /label= transmembrane\_region  
 FT 580..590  
 FT /label= ATP\_binding\_motif  
 FT 760..761  
 FT /label= autophosphorylation\_sites  
 FT /note= "putative"  
 FT 756..756  
 FT /label= autophosphorylation\_site  
 FT /note= "putative"  
 FT 802..805  
 FT /label= binding\_motif\_for\_P13\_kinase  
 FT /note= "binding motif for phosphatidylinositol 3' kinase"  
 FT 790  
 FT /label= potential\_substrate\_binding\_site  
 FT 26..42  
 FT /note= "antibody recognition sequence NTalpha"  
 FT 309..321  
 FT /note= "antibody recognition sequence NTBeta"  
 FT 860..877  
 FT /note= "antibody recognition sequence CTbeta"  
 US5677144-A.  
 14-OCT-1997.  
 08-NOV-1994; 336343.  
 16-NOV-1993; US-153397.  
 (ALVE/) ALVES F H E.  
 (ULLR/) ULLRICH A.  
 PI Alves FHE, Ullrich A;  
 DR WPI; 97-511869/47.  
 Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding  
 for it, useful for cancer diagnosis  
 PS Disclosure; Page -; 70pp; English.  
 The present sequence represents a splice variant of a mammary



carcinoma kinase (MCK-10). This kinase belongs to a novel family of receptor tyrosine kinases, and expression is associated with proliferative diseases such as cancer. The MCK-10 receptor tyrosine kinase has extensive sequence similarity to the insulin receptor family. The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide primer pools, using a template cDNA synthesised by reverse transcription of poly-A RNA from the human mammary carcinoma cell line MCF7. The amplified PCR product was used to screen human foetal brain and placental libraries, from which the present splice variant was isolated. This splice variant does not possess amino acids 505-541 of MCK-10 (W34672). The sequence represented by amino acids 548-558 may be important, as serine/threonine kinase results in reduced ligand binding affinity. MCK-10 is expressed in brain tissue, and the protein shares homology with the tyrosine kinase neurotrophin receptor. Modulation of MCK-10 activity therefore may be used for treatment of neurological disorders. MCK-10 is also expressed in a variety of cancer cell lines and tumour tissue. The nucleotide sequence of MCK-10, or parts of it, can be used for diagnostic purposes to detect aberrant expression of MCK-10 genes. Inhibitors of MCK-10 (or splice variants) receptor activity may have therapeutic value in the treatment of diseases such as cancer.

note: the present sequence does not appear in the specification, but was created using information provided.

Query Match 36.4%; Score 100; DB 1; Length 876;  
Best Local Similarity 51.9%; Pred. No. 7.24e+00;  
Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;

Db 418 ILIGCLVAIILLLLLIIIALMLRLHWR 444  
QY 233 ILISSL-AIILMVSLLL-LSLWKL-WR 256

RESULT 4

ID W34673 standard; Protein; 882 AA.  
AC W34673;  
DE Human mammary carcinoma kinase 10 (MCK-10) splice variant 1.  
KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;  
KW Proliferative disease; cancer; insulin receptor family;  
KW tyrosine kinase neurotrophin receptor; MCK-10 activity;  
KW neurological disorder; aberrant expression.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT Protein 19..919  
FT Domain 31..185  
FT Cleavage\_site 304..307  
FT Region 48..439  
FT Binding\_site 580..590  
FT Modified\_site 765..766  
FT Modified\_site 761  
FT Binding\_site 807..810  
FT Binding\_site 795..795  
FT Region 26..42  
FT Region 309..321

carcinoma kinase (MCK-10). This kinase belongs to a novel family of receptor tyrosine kinases, and expression is associated with proliferative diseases such as cancer. The MCK-10 receptor tyrosine kinase has extensive sequence similarity to the insulin receptor family. The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide primer pools, using a template cDNA synthesised by reverse transcription of poly-A RNA from the human mammary carcinoma cell line MCF7. The amplified PCR product was used to screen human foetal brain and placental libraries, from which the present splice variant was isolated. This splice variant does not possess amino acids 505-541 of MCK-10 (W34672). The sequence represented by amino acids 548-558 may be important, as serine/threonine kinase results in reduced ligand binding affinity. MCK-10 is expressed in brain tissue, and the protein shares homology with the tyrosine kinase neurotrophin receptor. Modulation of MCK-10 activity therefore may be used for treatment of neurological disorders. MCK-10 is also expressed in a variety of cancer cell lines and tumour tissue. The nucleotide sequence of MCK-10, or parts of it, can be used for diagnostic purposes to detect aberrant expression of MCK-10 genes. Inhibitors of MCK-10 (or splice variants) receptor activity may have therapeutic value in the treatment of diseases such as cancer.

note: the present sequence does not appear in the specification, but was created using information provided.

Query Match 36.4%; Score 100; DB 1; Length 876;  
Best Local Similarity 51.9%; Pred. No. 7.24e+00;  
Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;

Db 418 ILIGCLVAIILLLLLIIIALMLRLHWR 444  
QY 233 ILISSL-AIILMVSLLL-LSLWKL-WR 256

RESULT 5

ID R71100 standard; Protein; 914 AA.  
AC R71100; 1995 (first entry)  
DE Protein-tyrosine-kinase PTK22  
KW Protein-tyrosine-kinase; PTK; discoidin domain receptor; cancer;  
KW Breast tumor; mammary carcinoma; diagnosis; prognosis; therapy.  
OS Homo sapiens.  
PN W09502187-A.  
PF 19-JAN-1995.  
PF 08-JUL-1994; G01480.  
PF 09-JUL-1993; GB-014271.  
PR (CANC-) CANCER RES INST.  
PR (WELL) WELLCOME FOUND LTD.  
PR Barker KT, Crompton MR, Gusterson BA, Martindale JE;  
PI Mitchell PJ, Page MJ, Spence P;  
PI WPI: 95-066991/09.  
DR N-PSDB; 084782.  
DR Method for screening substances, using protein tyrosine kinase -  
PT for potential utility as therapeutic agents for cancer  
PT Disclosure; Page 26-30; 51pp; English.  
PS cDNA derived from tumor metastatic tissue was amplified using  
CC primers (given in 084783-84) based on sequences (R71101, R71103)  
CC associated with protein-tyrosine-kinases (PTK). Novel PTK22 was

/note= "antibody recognition sequence NTbeta"  
865..882  
/note= "antibody recognition sequence CTbeta"

Region  
US5677144-A.  
PN 14-OCT-1997.  
PD 08-NOV-1994; 336343.  
PR 16-NOV-1993; US-153397.  
PR (ALVE/) ALVES F H E.  
PA (ULR/) ULLRICH A.  
PI Alves FHE, Ullrich A;  
PI WPI: 97-511869/47.  
DR Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding  
PT for it, useful for cancer diagnosis  
PT Disclosure; Page: 70pp; English.  
PS The present sequence represents a splice variant of a mammary  
CC carcinoma kinase (MCK-10). This kinase belongs to a novel family  
CC of receptor tyrosine kinases, and expression is associated with  
CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine  
CC kinase has extensive sequence similarity to the insulin receptor family.  
CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide  
CC primer pools, using a template cDNA synthesised by reverse transcription  
CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The  
CC amplified PCR product was used to screen human foetal brain and placental  
CC libraries, from which the present splice variant was isolated. This  
CC splice variant does not possess amino acids 505-541 of MCK-10 (W34672).  
CC The sequence represented by amino acids 548-558 may be important, as  
CC deletion of this motif in the activin receptor serine/threonine kinase  
CC results in reduced ligand binding affinity. MCK-10 is expressed in brain  
CC tissue, and the protein shares homology with the tyrosine kinase  
CC neurotrophin receptor. Modulation of MCK-10 activity therefore may be used  
CC for treatment of neurological disorders. MCK-10 is also expressed in a  
CC variety of cancer cell lines and tumour tissue. The nucleotide sequence  
CC of MCK-10, or parts of it, can be used for diagnostic purposes to detect  
CC aberrant expression of MCK-10 genes. Inhibitors of MCK-10 (or splice  
CC variants) receptor activity may have therapeutic value in the treatment  
CC of diseases such as cancer. does not appear in the specification, but was  
CC note: the present sequence does not appear in the specification, but was  
CC created using information provided.  
SQ Sequence 882 AA;

CC identified in an isolated subclone. The 3' sequence of PTK22 was  
 CC obtained by reverse transcription (using the primer of Q84786) and  
 CC PCR amplification (primers Q84787-88) of RNA of human breast  
 CC carcinoma cell line MDA MG 468. The partial DNA sequence of PTK22  
 CC is given in Q84782.  
 CC Sequence 914 AA;

Query Match 36.4%; Score 100; DB 1; Length 914;  
 Best Local Similarity 51.9%; Pred. No. 7.24e+00;  
 Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;  
 Db 419 ILIGCLVAIIILLIIILALMLRWLR 445  
 QY 233 ILISL-AILLMVSLLL-LSLWKL-WR 256

RESULT 6  
 ID W34672 standard; Protein; 919 AA.  
 AC W34672;  
 DT 17-FEB-1998 (first entry)  
 DE Human mammary carcinoma kinase 10 (MCK-10) amino acid sequence.  
 KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;  
 KW proliferative disease; cancer; insulin receptor family;  
 KW tyrosine kinase neurotrophin receptor; MCK-10 activity;  
 KW neurological disorder; aberrant expression.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Peptide 1..18  
 FT /label= signal\_sequence  
 FT 19..919  
 FT /note= "mature\_protein"  
 FT Domain 31..185  
 FT /label= Discoidin\_I-like\_domain  
 FT Cleavage\_site 304..307  
 FT /label= endopeptidase\_furin  
 FT Region 48..439  
 FT /note= "putative precursor cleavage site"  
 FT Binding\_site 617..627  
 FT /label= transmembrane\_region  
 FT Modified\_site 802..803  
 FT /label= autophosphorylation\_sites  
 FT 798  
 FT /note= "putative"  
 FT Binding\_site 844..847  
 FT /label= autophosphorylation\_site  
 FT /note= "putative"  
 FT /label= binding\_motif\_for\_p13\_kinase  
 FT /note= "binding motif for phosphatidylinositol 3' kinase"  
 FT Binding\_site 832..832  
 FT /label= potential\_substrate\_binding\_site  
 FT Binding\_site 506..509  
 FT /label= putative\_receptor\_binding\_site\_for\_SHC  
 FT /note= "SHC is an oncogenic SH2 domain containing molecule"  
 FT Binding\_site 510..513  
 FT /label= GTPase\_activity\_protein\_binding\_site  
 FT /note= "putative"  
 FT Region 505..541  
 FT /note= "alternatively spliced variant"  
 FT Region 666..671  
 FT /note= "alternatively spliced sequence"  
 FT Region 26..42  
 FT /note= "antibody recognition sequence Nta1pha"  
 FT Region 309..321  
 FT /note= "antibody recognition sequence Ntbeta"  
 FT Region 902..919  
 FT /note= "antibody recognition sequence CTbeta"  
 US5677144-A.  
 PN 14-OCT-1997. 336343.  
 PD 08-NOV-1994; US-153397.  
 PR 16-NOV-1993; US-153397.

PA (ALVE/) ALVES F H E.  
 PA (ULIR/) ULLRICH A.  
 PI Alves FHE, Ullrich A;  
 DR WPI; 97-511869/47.  
 DR N-PSDB; T93785.  
 PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding  
 PT for it, useful for cancer diagnosis  
 PS Disclosure; Fig 1; 70pp; English.  
 CC The present sequence represents the protein sequence of a mammary  
 CC carcinoma kinase, called MCK-10. This kinase belongs to a novel family  
 CC of receptor tyrosine kinases, and expression is associated with  
 CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine  
 CC kinase has extensive sequence similarity to the insulin receptor family.  
 CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide  
 CC primer pools, using a template cDNA synthesised by reverse transcription  
 CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The MCK-10  
 CC protein contains 2 alternative spliced sequences, from amino acids  
 CC 505-541 and 666-671. The sequence represented by amino acids 585-595 may  
 CC be important, as deletion of this motif in the activin receptor  
 CC serine/threonine kinase results in reduced ligand binding affinity.  
 CC MCK-10 is expressed in brain tissue, and the protein shares homology with  
 CC the tyrosine kinase neurotrophin receptor. Modulation of MCK-10 activity  
 CC therefore may be used for treatment of neurological disorders. MCK-10 is  
 CC also expressed in a variety of cancer cell lines and tumour tissue. The  
 CC nucleotide sequence of MCK-10, or parts of it, can be used for diagnostic  
 CC purposes to detect aberrant expression of MCK-10 genes. Inhibitors of  
 CC MCK-10 receptor activity may have therapeutic value in the treatment of  
 CC diseases such as cancer.  
 CC Sequence 919 AA;

Query Match 36.4%; Score 100; DB 1; Length 919;  
 Best Local Similarity 51.9%; Pred. No. 7.24e+00;  
 Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;

Db 418 ILIGCLVAIIILLIIILALMLRWLR 444  
 QY 233 ILISL-AILLMVSLLL-LSLWKL-WR 256

RESULT 7  
 ID R75502 standard; Protein; 919 AA.  
 AC R75502;  
 DT 26-NOV-1995 (first entry)  
 DE Human mammary carcinoma kinase 10 (MCK-10).  
 KW Mammary carcinoma kinase 10; transmembrane receptor;  
 KW receptor tyrosine kinase; cancer.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Peptide 1..18  
 FT /label= signal  
 FT domain 31..185  
 FT /label= discoidin\_I-like\_domain  
 FT cleavage\_site 304..307  
 FT /label= putative\_precursor\_cleavage\_site  
 FT region 417..439  
 FT /label= transmembrane  
 FT misc\_difference 505..541  
 FT /label= alternatively\_spliced\_sequence\_I  
 FT misc\_difference 666..671  
 FT /label= alternatively\_spliced\_sequence\_II  
 FT misc\_difference 25..42  
 FT /label= NT\_alpha  
 FT /note= "peptide antibody recognition site"  
 FT misc\_difference 309..321  
 FT /label= NT\_beta  
 FT /note= "see above"  
 FT misc\_difference 909..919  
 FT /label= CT\_beta  
 FT /note= "see above"  
 WO9514088-A.  
 PN 26-MAY-1995.  
 PD 16-NOV-1994; E03797.  
 PR 16-NOV-1993; US-153397.





Thu May 11 06:49:39 2000

```

FT binding_site 324..332
FT /note= "Putative cation binding domain"
FT binding_site 386..394
FT /note= "Putative cation binding domain"
FT binding_site 441..449
FT /note= "Putative cation binding domain"
FT domain 1040..1044
FT /label= "Cytoplasmic domain"
FT /note= "Conserved in virtually all integrins"
FT misc_difference 1044..1045
FT /note= "Position of deletion of alpha 6A"
FT WO9219647-A.
FT 12-NOV-1992.
FT 27-APR-1992; U03527.
FT 03-MAY-1991; US-695564.
FT (SCRI) SCRIpps RES INST.
FT Quaranta V, Tamura RN;
FT WPI; 92-398789/48.
FT N-PSDB; Q3189.
FT Integrin alpha sub-unit cytoplasmic domain polypeptide(s) - used
FT for prodn. of antibodies and in detection of integrin sub-units
FT in body samples
FT Disclosure; Page 78-82; 115pp; English.
FT The sequences given in R28821-22 are the human alpha 6A and 6B
FT integrin subunits. Integrins are a family of cell surface receptors
FT which serve cellular adhesion functions. These receptors form a link
FT between the extracellular matrix and the cytoskeleton through their
FT binding to various extracellular components. Each integrin receptor
FT is a heterodimer comprised of an alpha and a beta subunit. Each alpha
FT subunit tends to associate with only one type of beta subunit but
FT there are several exceptions to this rule. These integrins correspond
FT to the laminin receptor. The cytoplasmic domain of the 6A and 6B
FT integrins differs from previously isolated alpha 6 integrins. The
FT human alpha 6B was isolated from human choriocarcinoma cell line JAR
FT by immunoprecipitation studies.
FT Sequence 1091 AA;

SQ Query Match 36.0%; Score 99; DB 1; Length 1091;
Best Local Similarity 50.0%; Pred. No. 8.39e+00;
Matches 11; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Db 1017 IILVAILAGILMLLVFLWK 1038
QY 232 FILISSLAILLMVSLLSLWK 253

RESULT 12
ID WO3761 standard; Protein; 505 AA.
AC WO3761; 1996 (first entry)
DE Mullerian inhibiting substance receptor MISR3.
KW Mullerian inhibiting substance receptor; MISR; TGF-beta receptor;
KW transforming growth factor beta type I receptor; gene therapy;
KW wound healing; tumour treatment; rat inhibin.
OS Rattus sp.
FH Key Location/Qualifiers
FT binding_site 210..230
FT /label= ATP_binding_site
FT /note= "corresponds to conserved GXGXXGVX(11-28)K
FT motif found in all serine/threonine kinases
FT and thought to form an ATP binding site"
FT
FT US5538892-A.
PN 23-JUL-1996.
PD 18-MAR-1992; 853396.
PR 18-MAR-1992; US-853396.
PR 11-MAR-1993; US-029673.
PR 04-NOV-1993; US-149105.
PR (GHEO) GEN HOSPITAL CORP.
PA (YDUU) UNIV DUKE.
PI Donahoe PK, Gustafson M, He W, Wang X;
DR N-PSDB; T36071.
PT New isolated TGF-beta type I receptor DNA - used to develop prods

```

```

PT for diagnosis and therapy, e.g. for treating tumours or promoting
PT wound healing
PS Disclosure; Columns 45-48; 44pp; English.
PS Degenerate PCR primers were designed based on two highly conserved
CC regions within the cDNA encoding a murine activin receptor, human
CC and porcine TGF-beta type II receptor and the daf-1 receptor of
CC C.elegans. The primers (see T36072 and T36073) were used for
CC amplifying clones present in a 14.5 day foetal rat urogenital ridge
CC cDNA COS cell expression library. Four clones encoding portions of
CC four novel polypeptides (all putative serine/threonine kinases)
CC were obtained and inserts from these clones were used as probes to
CC isolate full-length cDNA sequences for each of the four TGF-beta
CC type I receptors. Each putative receptor comprises a hydrophobic
CC signal peptide of 19-23 residues, an extracellular, cysteine-rich,
CC hydrophilic, ligand-binding domain of 100-150 residues, a
CC hydrophobic single transmembrane domain of 23-25 residues, an
CC intracellular serine/threonine kinase domain of approximately 300
CC residues and a short serine/threonine-rich tail. The present sequence
CC is that of MISR3 which is believed to be a monomeric isoform of the
CC rat inhibin receptor and/or BMP receptor.
CC isolated DNA which encodes the MISR4 sequence or which is able to
CC hybridise to such DNA under stringent conditions is claimed.
SQ Sequence 505 AA;

Query Match 34.9%; Score 96; DB 1; Length 505;
Best Local Similarity 44.0%; Pred. No. 1.30e+01;
Matches 11; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

Db 122 LILGPVALLVVALGTGLWVR 146
QY 232 FILISSLAILLMVSLLSLWK 256

RESULT 13
ID R41922 standard; Protein; 505 AA.
AC R41922;
DE 20-APR-1994 (first entry)
DE MISR3
KW MIS; Mullerian Inhibitory Substance; receptor;
KW transforming growth factor; inhibin; BMP;
KW membrane serine/threonine kinase receptor;
KW bone morphogenesis protein.
OS Rattus rattus.
PN WO9319177-A.
PD 30-SEP-1993.
PF 15-MAR-1993; U02387.
PF 18-MAR-1992; US-853396.
PR 11-MAR-1993; US-029673.
PR (GHEO) GEN HOSPITAL CORP.
PI Donahoe PK, Gustafson M, He W;
DR WPI; 93-320743/40.
DR N-PSDB; Q49765.
DR New receptors of the transforming growth factor-beta receptor
PT family - comprising Mullerian Inhibitory Substance Receptors and
PT inhibin receptors
PS Claim 51; Fig 3; 59pp; English.
CC Misr1 (Q49763) is believed to encode an isoform of the rat
CC MIS receptor. Misr2A/misr2B (Q49764), misr3 (Q49765) and misr4
CC (Q49766) are believed to encode monomeric isoforms of the rat
CC inhibin receptor and/or BMP receptor.
CC Sequence 505 AA;

Query Match 34.9%; Score 96; DB 1; Length 505;
Best Local Similarity 44.0%; Pred. No. 1.30e+01;
Matches 11; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

Db 122 LILGPVALLVVALGTGLWVR 146
QY 232 FILISSLAILLMVSLLSLWK 256

RESULT 14
ID R26322 standard; Protein; 788 AA.

```

AC R26322;  
 DT 29-JAN-1993 (first entry)  
 DE B6H.  
 KW Beta-6; beta-1; beta-2; beta-3; cytoplasmic tail; ligand; vitronectin;  
 KW epithelial-derived; tumour cells; fibronectin.  
 OS Homo sapiens.  
 PN W09212236-A.  
 PD 23-JUL-1992.  
 PE 11-JAN-1991; U00236.  
 PR 11-JAN-1991; WO-U02362.  
 PA (SCRI) SCRIPPS CLINIC & RES CENT.  
 PA (REGC) UNIV CALIFORNIA.  
 PI Pytela R, Quaranta V, Sheppard D;  
 DR WPI; 92-284332/34.  
 DR N-PSDB; Q27642.  
 PT New integrin beta sub-unit and its nucleic acid - forms  
 PT hetero-dimers with sub-units alpha-V and alpha-E, useful as a  
 PT diagnostic  
 PS Claim 2; Fig 3; 43pp; English.  
 CC The sequences given in R26322-23 are human and guinea pig integrin beta  
 CC subunit, beta-6 proteins. The beta-6 subunit is a surface receptor  
 CC which is useful in mediating critical aspects of cell processes in  
 CC conjunction with an integrin alpha subunit. Beta-6 is clearly  
 CC different from beta-1, beta-2, beta-3 and other beta subunits which  
 CC have been discovered. The 11 amino acid carboxyl-terminal extension  
 CC distinguishes it from other beta subunits. This cytoplasmic tail  
 CC indicates that beta-6 regulates signal transduction differently from  
 CC other beta subunits. Beta-6 has been isolated from epithelial-derived  
 CC tumour cells in association with the integrin alpha subunit, alpha-V.  
 CC Beta-6 can associate with a variety of alpha subunits to form a  
 CC functional integrin. Alpha-V/beta-6 can bind extracellular matrix  
 CC molecules, eg. ligands, pref. containing the sequence Arg-Gly-Asp,  
 CC such as vitronectin or fibronectin.  
 SQ Sequence 788 AA;

Query Match 34.9%; Score 96; DB 1; Length 788;  
 Best Local Similarity 42.9%; Pred. No. 1.30e+01;  
 Matches 12; Conservative 11; Mismatches 4; Indels 1; Gaps 1;

Db 706 PNIPMIMGVSLATLL-IGVLLCIWKL 732  
 QY 227 PKLSKFISSLAILLMVSLLSLWKL 254

## RESULT 15

ID W02194 standard; Protein; 799 AA.  
 AC W02194;  
 DT 08-OCT-1996 (first entry)  
 DE Human integrin beta subunit protein, beta-5.  
 KW Human; integrin beta subunit; beta-5; carcinoma; lymphoid cell;  
 KW immunoassay; detection; mRNA; assay.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..23  
 FT /label= sig\_peptide  
 FT peptide 24..799  
 FT /label= mat\_peptide  
 FT domain 720..742  
 FT /note= "transmembrane domain"  
 FT region 460..462  
 FT /note= "glycosylation site"  
 FT region 477..479  
 FT /note= "glycosylation site"  
 FT region 505..507  
 FT /note= "glycosylation site"  
 FT region 552..554  
 FT /note= "glycosylation site"  
 FT region 586..588  
 FT /note= "glycosylation site"  
 FT region 654..656  
 FT /note= "glycosylation site"  
 FT region 705..707  
 FT /note= "glycosylation site"

FT region 791..793  
 FT /note= "glycosylation site"  
 FT region 795..797  
 FT /note= "glycosylation site"  
 PN US5527679-A.  
 PD 18-JUN-1996.  
 PE 01-MAY-1991; 694314.  
 PR 01-MAY-1991; US-694314.  
 PR 27-APR-1993; US-054077.  
 PA (DAND) DANA FARBER CANCER INST INC.  
 PI Hemler ME, Ramaswamy H;  
 DR WPI; 96-299852/30.  
 DR N-PSDB; T36481.  
 PT Immunoassay and mRNA hybridisation assay for beta-5 protein - useful  
 PT for the detection of carcinoma(s) and to distinguish different cell  
 PT types  
 PS Claim 12; Columns 13-20; 21pp; English.  
 CC The present sequence is the human integrin beta subunit  
 CC protein, beta-5, which is found in carcinomas but not in lymphoid  
 CC cells. An immunoassay for the detection of beta-5, comprises  
 CC contacting a sample with a monoclonal antibody (Ab) which binds 1  
 CC epitope of beta-5, and then with a labelled Ab which binds another  
 CC epitope of beta-5, and detecting any bound label. An assay for  
 CC beta-5 mRNA, comprises contacting a sample with a probe capable of  
 CC hybridising to the beta-5 cDNA, and determining if binding has  
 CC occurred. These assays are useful for detecting carcinomas, and for  
 CC distinguishing between different cell types.  
 SQ Sequence 799 AA;  
 Query Match 34.9%; Score 96; DB 1; Length 799;  
 Best Local Similarity 45.5%; Pred. No. 1.30e+01;  
 Matches 10; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Db 723 ILLAVVGSILLVGLALLATWKL 744  
 QY 233 ILISSLAILLMVSLLSLWKL 254

Search completed: Wed May 10 12:01:23 2000  
 Job time : 8 secs.

US-09-376-430-2-07.rai

Thu May 11 06:49:39 2000

\*\*\*\*\*  
M P S R L H  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 12:01:40 2000; Maspar time 44.65 Seconds  
10.155 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-376-430-2  
Description: (226-260) from US09376430A.pep (7 of 25)  
Perfect Score: 275  
Sequence: 1 KPGLSFILLISLAILLMVSLLLSLWKLWRVKKF 35

Scoring table: PAM 150  
Gap 11

Searched: 131253 seqs, 12956647 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-issued  
1:5A\_COMB 2:5B\_COMB 3:PCT\_COMB 4:backfiles1

Statistics: Mean 23.111; Variance 132.169; scale 0.175

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	100	36.4	913	1	US-08-445-640-4	Sequence 4, Applicatio
2	100	36.4	913	1	US-08-445-640-4	Sequence 2, Applicatio
3	99	36.0	1073	1	US-08-445-640-4	Sequence 1, Applicatio
4	99	36.0	1073	1	US-08-445-640-4	Sequence 1, Applicatio
5	99	36.0	1091	1	US-08-445-640-4	Sequence 3, Applicatio
6	99	36.0	1091	1	US-08-445-640-4	Sequence 3, Applicatio
7	96	34.9	505	1	US-08-445-640-4	Sequence 16, Applicati
8	96	34.9	505	1	US-08-445-640-4	Sequence 16, Applicati
9	96	34.9	788	2	US-08-445-640-4	Sequence 27, Applicati
10	96	34.9	799	1	US-08-445-640-4	Sequence 2, Applicatio
11	95	34.5	141	1	US-08-445-640-4	Sequence 5, Applicatio
12	95	34.5	141	1	US-08-445-640-4	Sequence 5, Applicatio
13	95	34.5	149	1	US-08-445-640-4	Sequence 7, Applicatio
14	95	34.5	149	1	US-08-445-640-4	Sequence 7, Applicatio
15	95	34.5	1039	4	5196511-2	Patent No. 5196511
16	94	34.2	1285	3	PCT-US96-1	Sequence 6, Applicatio
17	94	34.2	1285	3	PCT-US96-1	Sequence 6, Applicatio
18	91	33.1	212	2	US-08-445-640-4	Sequence 4, Applicatio
19	91	33.1	470	2	US-08-445-640-4	Sequence 7, Applicatio
20	91	33.1	470	2	US-08-445-640-4	Sequence 7, Applicatio
21	91	33.1	471	2	US-08-445-640-4	Sequence 7, Applicatio
22	91	33.1	471	2	US-08-445-640-4	Sequence 7, Applicatio
23	91	33.1	471	2	US-08-445-640-4	Sequence 2, Applicatio

24	91	33.1	471	2	US-08-394-	Sequence 2, Applicatio	1.38e+01
25	91	33.1	476	1	US-08-737-	Sequence 1, Applicatio	1.38e+01
26	91	33.1	498	1	US-08-416-	Sequence 9, Applicatio	1.38e+01
27	91	33.1	498	2	US-08-474-	Sequence 9, Applicatio	1.38e+01
28	91	33.1	498	2	US-08-394-	Sequence 9, Applicatio	1.38e+01
29	91	33.1	503	2	US-08-696-	Sequence 2, Applicatio	1.38e+01
30	91	33.1	503	3	PCT-US95-0	Sequence 2, Applicatio	1.38e+01
31	91	33.1	503	3	PCT-US94-1	Sequence 4, Applicatio	1.38e+01
32	91	33.1	503	2	US-08-481-	Sequence 2, Applicatio	1.38e+01
33	90	32.7	401	4	5252556-1	Patent No. 5252556	1.59e+01
34	89	32.4	130	3	PCT-US96-1	Sequence 20, Applicati	1.84e+01
35	89	32.4	130	2	US-08-494-	Sequence 16, Applicati	1.84e+01
36	89	32.4	202	2	US-08-494-	Sequence 16, Applicati	1.84e+01
37	89	32.4	202	3	PCT-US96-1	Sequence 4, Applicatio	2.12e+01
38	88	32.0	860	1	US-08-093-	Sequence 16, Applicati	2.44e+01
39	87	31.6	774	1	US-08-070-	Sequence 50, Applicati	2.44e+01
40	87	31.3	989	2	US-08-419-	Sequence 13, Applicati	2.82e+01
41	86	31.3	91	2	PCT-US92-0	Sequence 15, Applicati	2.82e+01
42	86	31.3	91	1	US-07-847-	Sequence 13, Applicati	2.82e+01
43	86	31.3	91	2	US-08-440-	Sequence 13, Applicati	2.82e+01
44	86	31.3	91	2	US-08-436-	Sequence 15, Applicati	2.82e+01
45	86	31.3	91	2	US-08-436-	Sequence 15, Applicati	2.82e+01

ALIGNMENTS

RESULT 1  
ID US-08-445-640-4 STANDARD; PRT: 913 AA.

XX xxxxxx

XX Sequence 4, Application US/08445640

DE Sequence 4, Application US/08445640

XX Patent No. 5709458

CC GENERAL INFORMATION:

CC APPLICANT: Godowski, Paul J.

CC APPLICANT: Mark, Melanie R.

CC APPLICANT: Scadden, David T.

CC APPLICANT: Baker, Kevin P.

CC APPLICANT: Baron, Will F.

CC TITLE OF INVENTION: Protein Tyrosine Kinases

CC NUMBER OF SEQUENCES: 35

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Genentech, Inc.

CC STREET: 460 Point San Bruno Blvd

CC CITY: South San Francisco

CC STATE: California

CC COUNTRY: USA

CC ZIP: 94080

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patin (Genentech)

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/445,640

CC FILING DATE: 22-MAY-1995

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: 08/170558

CC FILING DATE: 20-DEC-1993

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: 08/157563

CC FILING DATE: 23-NOV-1993

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Hasak, Janet E.

CC REGISTRATION NUMBER: 28,616

CC REFERENCE/DOCKET NUMBER: 854C2

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 415/225-1896

CC TELEFAX: 415/952-9881  
CC TELEX: 910/371-7168  
CC INFORMATION FOR SEQ ID NO: 4:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 913 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC SEQUENCE 913 AA; 101153 MW; 4161835 CN;

Query Match 36.4%; Score 100; DB 1; Length 913;  
Best Local Similarity 51.9%; Pred. No. 3.74e+00;  
Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;

Db 418 ILIGCLVAIIILLIIILMLWRLHWR 444  
QY 233 ILISSL-AIILMVSLLL-LSLWKL-WR 256

RESULT 2  
ID US-08-336-343A-2 STANDARD; PRT: 919 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX

Sequence 2, Application US/08336343A  
DE Sequence 2, Application US/08336343A  
XX Patent No. 5677144  
CC GENERAL INFORMATION:  
CC APPLICANT: Ullrich, Axel

CC APPLICANT: Alves, Frauke  
CC TITLE OF INVENTION: CCR-2, A No. 5677144el Receptor Tyrosine Kinase  
CC NUMBER OF SEQUENCES: 43  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: U.S.A.  
CC ZIP: 10036-2711

COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/336,343A  
CC FILING DATE: 08-NOV-1994  
CC CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
CC NAME: Coruzzi, Laura A.  
CC REGISTRATION NUMBER: 30,742  
CC REFERENCE/DOCKET NUMBER: 7683-065  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 790-9090  
CC TELEX: (212) 869-9741/8864  
CC TELEFAX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 919 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
CC SEQUENCE 919 AA; 101795 MW; 4228416 CN;

Query Match 36.4%; Score 100; DB 1; Length 919;  
Best Local Similarity 51.9%; Pred. No. 3.74e+00;  
Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;

Db 418 ILIGCLVAIIILLIIILMLWRLHWR 444  
QY 233 ILISSL-AIILMVSLLL-LSLWKL-WR 256

RESULT 3  
ID US-08-241-387-1 STANDARD; PRT: 1073 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX

Sequence 1, Application US/08241387  
DE Sequence 1, Application US/08241387  
XX Patent No. 5589570  
CC GENERAL INFORMATION:  
CC APPLICANT: Tamura, Richard N.

CC APPLICANT: Quaranta, Vito  
CC TITLE OF INVENTION: INTEGRIN ALPHA SUBUNIT CYTOPLASMIC  
CC TITLE OF INVENTION: DOMAIN POLYPEPTIDES, ANTIBODIES AND METHODS  
CC NUMBER OF SEQUENCES: 16  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: The Scripps Research Institute  
CC STREET: 10666 No. 5589570th Torrey Pines Road, TPC-8  
CC CITY: La Jolla  
CC STATE: CA  
CC COUNTRY: US  
CC ZIP: 92037

COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/241,387  
CC FILING DATE: 10-MAY-1994  
CC CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: USSN 07/695,564  
CC FILING DATE: 03-MAY-1004  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Fitting, Thomas  
CC REGISTRATION NUMBER: 34,163  
CC REFERENCE/DOCKET NUMBER: TSRI241.0D1

TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 619-554-2937  
CC TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 1:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1073 amino acids  
CC TYPE: amino acid

TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
CC HYPOTHETICAL: NO  
CC ANTI-SENSE: NO  
CC ORIGINAL SOURCE:  
CC ORGANISM: Homo sapiens

FEATURE:  
CC NAME/KEY: Domain  
CC LOCATION: 1012..1037  
CC OTHER INFORMATION: /label= TRANSMEMBRANE  
CC OTHER INFORMATION: /note= "The putative transmembrane region is  
CC OTHER INFORMATION: encompassed by amino acids 1012-1037."

FEATURE:  
CC NAME/KEY: Cleavage-site  
CC LOCATION: (23-24)  
CC OTHER INFORMATION: /note= "The mature protein is  
CC OTHER INFORMATION: cleaved from the signal sequence between amino  
CC OTHER INFORMATION: acids 23-24."

FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 223  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."



CC NAME/KEY: Modified-site  
CC LOCATION: 284  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 370  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 513  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 731  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 748  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 891  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 927  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 958  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Binding-site  
CC LOCATION: 230..238  
CC OTHER INFORMATION: /note= "Represents a putative  
CC OTHER INFORMATION: cation binding domain."  
CC FEATURE:  
CC NAME/KEY: Binding-site  
CC LOCATION: 324..332  
CC OTHER INFORMATION: /note= "Represents a putative  
CC OTHER INFORMATION: cation binding domain."  
CC FEATURE:  
CC NAME/KEY: Binding-site  
CC LOCATION: 386..394  
CC OTHER INFORMATION: /note= "Represents a putative  
CC OTHER INFORMATION: cation binding domain."  
CC FEATURE:  
CC NAME/KEY: Binding-site  
CC LOCATION: 441..449  
CC OTHER INFORMATION: /note= "Represents a putative  
CC OTHER INFORMATION: cation binding domain."  
CC FEATURE:  
CC NAME/KEY: Domain  
CC LOCATION: 1040..1044  
CC OTHER INFORMATION: /label= CYTOPLASMIC  
CC OTHER INFORMATION: /note= "The cytoplasmic sequence, which is  
CC OTHER INFORMATION: conserved in virtually all of the integrin ALPHA  
CC OTHER INFORMATION: chains."  
CC FEATURE:  
CC NAME/KEY: Region  
CC LOCATION: 927..1073  
CC OTHER INFORMATION: /note= "The sequence encoded by the  
CC OTHER INFORMATION: fragment of ALPHA 6A cDNA amplified using primers  
CC OTHER INFORMATION: 1156/1157."  
CC SEQUENCE 1073 AA; 119575 MW; 5953577 CN;

Query Match 36.0%; Score 99; DB 1; Length 1073;  
Best Local Similarity 50.0%; Pred. No. 4.33e+00;  
Matches 11; Conservative 8; Mismatches 3; Indels 0; Gaps 0;  
Db 1017 IILVAILGILMLALLVFLWK 1038  
QY 232 FILISSLAILLMSVLSLLSLWK 253  
RESULT 4  
ID US-07-695-564-1 STANDARD; PRT; 1073 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE  
XX  
CC Sequence 1, Application US/07695564  
CC Patent No. 5310874  
CC GENERAL INFORMATION:  
CC APPLICANT: Tamura, Richard N.  
CC APPLICANT: Quaranta, Vito  
CC TITLE OF INVENTION: INTEGRIN ALPHA SUBUNIT CYTOPLASMIC  
CC TITLE OF INVENTION: DOMAIN POLYPEPTIDES, ANTIBODIES AND METHODS  
CC NUMBER OF SEQUENCES: 16  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Thomas Fitting  
CC STREET: 11300 Sorrento Valley Road, Suite 200  
CC CITY: San Diego  
CC STATE: California  
CC COUNTRY: United States  
CC ZIP: 92121  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent in Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA: US/07/695,564  
CC APPLICATION NUMBER:  
CC FILING DATE: 19910503  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER:  
CC FILING DATE:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Fitting, Thomas  
CC REGISTRATION NUMBER: 34,163  
CC REFERENCE/DOCKET NUMBER: SCRO377P  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 619-546-1555  
CC INFORMATION FOR SEQ ID NO: 1:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1073 amino acids  
CC TYPE: AMINO ACID  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
CC HYPOTHETICAL: NO  
CC ANTI-SENSE: NO  
CC ORIGINAL SOURCE:  
CC ORGANISM: Homo sapiens  
CC FEATURE:  
CC NAME/KEY: Domain  
CC LOCATION: 1012..1037  
CC OTHER INFORMATION: /label= TRANSMEMBRANE  
CC OTHER INFORMATION: /note= "The putative transmembrane region is  
CC OTHER INFORMATION: encompassed by amino acids 1012-1037."  
CC FEATURE:  
CC NAME/KEY: Cleavage-site  
CC LOCATION: (23,24)  
CC OTHER INFORMATION: /note= "The mature protein is  
CC OTHER INFORMATION: cleaved from the signal sequence between amino

CC OTHER INFORMATION: acids 23-24."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 223  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 284  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 370  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 513  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 731  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 748  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 891  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 927  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Modified-site  
CC LOCATION: 958  
CC OTHER INFORMATION: /label= GLYCOSYLATION  
CC OTHER INFORMATION: /note= "Potential site of N-linked glycosylation."  
CC FEATURE:  
CC NAME/KEY: Binding-site  
CC LOCATION: 230..238  
CC OTHER INFORMATION: /note= "Represents a putative  
CC OTHER INFORMATION: cation binding domain."  
CC FEATURE:  
CC NAME/KEY: Binding-site  
CC LOCATION: 324..332  
CC OTHER INFORMATION: /note= "Represents a putative  
CC OTHER INFORMATION: cation binding domain."  
CC FEATURE:  
CC NAME/KEY: Binding-site  
CC LOCATION: 385..394  
CC OTHER INFORMATION: /note= "Represents a putative  
CC OTHER INFORMATION: cation binding domain."  
CC FEATURE:  
CC NAME/KEY: Binding-site  
CC LOCATION: 441..449  
CC OTHER INFORMATION: /note= "Represents a putative  
CC OTHER INFORMATION: cation binding domain."  
CC FEATURE:  
CC NAME/KEY: Domain  
CC LOCATION: 1040..1044  
CC OTHER INFORMATION: /label= CYTOPLASMIC  
CC OTHER INFORMATION: /note= "The cytoplasmic sequence, which is  
CC OTHER INFORMATION: conserved in virtually all of the integrin ALPHA  
CC OTHER INFORMATION: chains."

CC CC  
CC NAME/KEY: Region  
CC LOCATION: 927..1073  
CC OTHER INFORMATION: /note= "The sequence encoded by the  
CC OTHER INFORMATION: fragment of ALPHA 6A cDNA amplified using primers  
CC OTHER INFORMATION: 1156/1157."  
CC SQ SEQUENCE 1073 AA; 119575 MW; 5953577 CN;  
Query Match 36.0%; Score 99; DB 1; Length 1073;  
Best Local Similarity 50.0%; Pred. No. 4.33e+00;  
Matches 11; Conservative 8; Mismatches 3; Indels 0; Gaps 0;  
Db 1017 ILVAILAGILMLALLVFLWK 1038  
QY 232 FILSSLAILLMVSLLSLWK 253  
RESULT 5  
ID US-08-241-387-3 STANDARD; PRT; 1091 AA.  
XX XX  
AC xxxxxx  
XX XX  
DT DT  
XX XX  
DE DE  
XX XX  
Sequence 3, Application US/08241387  
CC Sequence 3, Application US/08241387  
CC Patent No. 5589570  
CC GENERAL INFORMATION:  
CC APPLICANT: Tamura, Richard N.  
CC APPLICANT: Ouaranta, Vito  
CC TITLE OF INVENTION: INTEGRIN ALPHA SUBUNIT CYTOPLASMIC  
CC TITLE OF INVENTION: DOMAIN POLYPEPTIDES, ANTIBODIES AND METHODS  
CC NUMBER OF SEQUENCES: 16  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: The Scripps Research Institute  
CC STREET: 10666 No. 5589570th Torrey Pines Road, TPC-8  
CC CITY: La Jolla  
CC STATE: CA  
CC COUNTRY: US  
CC ZIP: 92037  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/241,387  
CC FILING DATE: 10-MAY-1994  
CC CLASSIFICATION: 530  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: USSN 07/695,564  
CC FILING DATE: 03-MAY-1004  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Fitting, Thomas  
CC REGISTRATION NUMBER: 34,163  
CC REFERENCE/DOCKET NUMBER: TSRI241.001  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 619-554-2937  
CC TELEFAX: 619-554-6312  
CC INFORMATION FOR SEQ ID NO: 3:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1091 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
CC HYPOTHEICAL: NO  
CC ANTI-SENSE: NO  
CC ORIGINAL SOURCE:  
CC ORGANISM: Homo sapiens  
CC FEATURE:  
CC NAME/KEY: Region  
CC LOCATION: 1..1091

CC OTHER INFORMATION: /note= "SEQ ID NO:3 is the 1091  
CC OTHER INFORMATION: residue amino acid sequence of the human ALPHA 6B  
CC OTHER INFORMATION: protein."  
CC FEATURE:  
CC NAME/KEY: Region  
CC LOCATION: 1..1044  
CC OTHER INFORMATION: /note= "The sequence of SEQ ID NO:3  
CC OTHER INFORMATION: is identical to SEQ ID NO:1 between amino acids 1  
CC OTHER INFORMATION: and 1044."  
CC FEATURE:  
CC NAME/KEY: Region  
CC LOCATION: 927..1060  
CC OTHER INFORMATION: /note= "Encompasses the sequence  
CC OTHER INFORMATION: encoded by the fragment of ALPHA 6B cDNA amplified  
CC OTHER INFORMATION: using primers 1156/1157."  
CC SEQUENCE 1091 AA; 122179 MW; 6176206 CN;  
SQ

Query Match 36.0%; Score 99; DB 1; Length 1091;  
Best Local Similarity 50.0%; Pred. No. 4.33e+00;  
Matches 11; Conservative 8; Mismatches 3; Indels 0; Gaps 0;  
DB 1017 IILVAILGILMLALLVFLWK 1038  
QY 232 FILISSLAILLVSVLLLSLWK 253

RESULT 6 STANDARD; PRT; 1091 AA.  
XX US-07-695-564-3  
AC xxxxxx  
DE Sequence 3, Application US/07695564  
XX Sequence 3, Application US/07695564  
CC Patent No. 5310874  
CC GENERAL INFORMATION:  
CC APPLICANT: Tamura, Richard N.  
CC APPLICANT: Quaranta, Vito  
CC TITLE OF INVENTION: INTEGRIN ALPHA SUBUNIT CYTOPLASMIC  
CC TITLE OF INVENTION: DOMAIN POLYPEPTIDES, ANTIBODIES AND METHODS  
CC NUMBER OF SEQUENCES: 16  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Thomas Fitting  
CC STREET: 11300 Sorrento Valley Road, Suite 200  
CC CITY: San Diego  
CC STATE: California  
CC COUNTRY: United States  
CC ZIP: 92121  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/07/695,564  
CC FILING DATE: 19910503  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER:  
CC FILING DATE:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Fitting, Thomas  
CC REGISTRATION NUMBER: 34,163  
CC REFERENCE/DOCKET NUMBER: SCRO377P  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 619-546-1555  
CC INFORMATION FOR SEQ ID NO: 3:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1091 amino acids  
CC TYPE: AMINO ACID  
CC TOPOLOGY: unknown

CC MOLECULE TYPE: protein  
CC HYPOTHETICAL: NO  
CC ANTI-SENSE: NO  
CC ORIGINAL SOURCE:  
CC ORGANISM: Homo sapiens  
CC FEATURE:  
CC NAME/KEY: Region  
CC LOCATION: 1..1091  
CC OTHER INFORMATION: /note= "SEQ ID NO:3 is the 1091  
CC OTHER INFORMATION: residue amino acid sequence of the human ALPHA 6B  
CC OTHER INFORMATION: protein."  
CC FEATURE:  
CC NAME/KEY: Region  
CC LOCATION: 1..1044  
CC OTHER INFORMATION: /note= "The sequence of SEQ ID NO:3  
CC OTHER INFORMATION: is identical to SEQ ID NO:1 between amino acids 1  
CC OTHER INFORMATION: and 1044."  
CC FEATURE:  
CC NAME/KEY: Region  
CC LOCATION: 927..1060  
CC OTHER INFORMATION: /note= "Encompasses the sequence  
CC OTHER INFORMATION: encoded by the fragment of ALPHA 6B cDNA amplified  
CC OTHER INFORMATION: using primers 1156/1157."  
CC SEQUENCE 1091 AA; 122179 MW; 6176206 CN;  
SQ

Query Match 36.0%; Score 99; DB 1; Length 1091;  
Best Local Similarity 50.0%; Pred. No. 4.33e+00;  
Matches 11; Conservative 8; Mismatches 3; Indels 0; Gaps 0;  
DB 1017 IILVAILGILMLALLVFLWK 1038  
QY 232 FILISSLAILLVSVLLLSLWK 253

RESULT 7 STANDARD; PRT; 505 AA.  
XX US-08-149-105-16  
AC xxxxxx  
DE Sequence 16, Application US/08149105  
XX Sequence 16, Application US/08149105  
CC Patent No. 5538892  
CC GENERAL INFORMATION:  
CC APPLICANT: Donahoe, Patricia K.  
CC APPLICANT: Gustafson, Michael  
CC APPLICANT: He, Wei W.  
CC APPLICANT: Wang, Xiao-Fan  
CC TITLE OF INVENTION: TGF-  
CC NUMBER OF SEQUENCES: 17  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Fish & Richardson  
CC STREET: 225 Franklin Street  
CC CITY: Boston  
CC STATE: Massachusetts  
CC COUNTRY: U.S.A.  
CC ZIP: 02110-2804  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
CC COMPUTER: IBM PS/2 Model 502 or 55sx  
CC OPERATING SYSTEM: MS-DOS (Version 5.0)  
CC SOFTWARE: WordPerfect (Version 5.1)  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/149,105  
CC FILING DATE:  
CC CLASSIFICATION: 530  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 08/029,673  
CC FILING DATE: March 11, 1993  
CC APPLICATION NUMBER: 07/853,396  
CC FILING DATE: March 18, 1992

CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Clark, Paul T.  
CC REGISTRATION NUMBER: 30,162  
CC REFERENCE/DOCKET NUMBER: 00786/211001  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (617) 542-5070  
CC TELEFAX: (617) 542-8906  
CC TELEX: 200154  
CC INFORMATION FOR SEQ ID NO: 16:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 505  
CC TYPE: amino acid  
CC STRANDEDNESS: linear  
CC TOPOLOGY: linear  
CC SEQUENCE 505 AA; 56774 MW; 1307601 CN;  
Query Match 34.9%; Score 96; DB 1; Length 505;  
Best Local Similarity 44.0%; Pred. No. 6.70e+00;  
Matches 11; Conservative 10; Mismatches 4; Indels 0; Gaps 0;  
Db 122 LILGPVLALLVLVALGTGLRWRR 146  
QY 232 FILISSLAILLMVSLLLSLWKLWR 256  
RESULT 8  
ID US-08-317-847-16 STANDARD; PRT; 505 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE  
XX  
Sequence 16, Application US/08317847  
Sequence 16, Application US/08317847  
Patent No. 5547854  
GENERAL INFORMATION:  
CC APPLICANT: Donahoe, Patricia K.  
CC APPLICANT: Gustafson, Michael  
CC APPLICANT: He, Wei W.  
CC TITLE OF INVENTION: FOUR NOVEL RECEPTORS OF THE TGF-B  
CC TITLE OF INVENTION: FAMILY  
CC NUMBER OF SEQUENCES: 17  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Fish & Richardson  
CC STREET: 225 Franklin Street  
CC CITY: Boston  
CC STATE: Massachusetts  
CC COUNTRY: U.S.A.  
CC ZIP: 02110-2804  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
CC COMPUTER: IBM PS/2 Model 50z or 55sx  
CC OPERATING SYSTEM: MS-DOS (Version 5.0)  
CC SOFTWARE: WordPerfect (Version 5.1)  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/317,847  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 08/029,673  
CC FILING DATE: March 11, 1993  
CC REFERENCE/DOCKET NUMBER: 00786/127002  
CC APPLICATION NUMBER: 07/853,396  
CC FILING DATE: March 18, 1992  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Clark, Paul T.  
CC REGISTRATION NUMBER: 30,162  
CC REFERENCE/DOCKET NUMBER: 00786/127002  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (617) 542-5070  
CC TELEFAX: (617) 542-8906  
CC TELEX: 200154  
CC INFORMATION FOR SEQ ID NO: 16:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 505  
CC TYPE: amino acid  
CC STRANDEDNESS: linear  
CC TOPOLOGY: linear  
CC SEQUENCE 505 AA; 56774 MW; 1307601 CN;  
Query Match 34.9%; Score 96; DB 1; Length 505;  
Best Local Similarity 44.0%; Pred. No. 6.70e+00;  
Matches 11; Conservative 10; Mismatches 4; Indels 0; Gaps 0;  
Db 122 LILGPVLALLVLVALGTGLRWRR 146  
QY 232 FILISSLAILLMVSLLLSLWKLWR 256  
RESULT 9  
ID US-07-728-215-27 STANDARD; PRT; 788 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE  
XX  
Sequence 27, Application US/07728215  
Sequence 27, Application US/07728215  
Patent No. 5962643  
GENERAL INFORMATION:  
CC APPLICANT: Sheppard, Dean  
CC APPLICANT: Quaranta, Vito  
CC APPLICANT: Pytela, Robert  
CC TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses  
CC TITLE OF INVENTION: Thereof  
CC NUMBER OF SEQUENCES: 43  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
CC STREET: 4370 La Jolla Village Drive, Suite 700  
CC CITY: San Diego  
CC STATE: California  
CC COUNTRY: United States of America  
CC ZIP: 92122  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/07/728,215  
CC FILING DATE: 19910711  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Campbell, Cathryn A.  
CC REGISTRATION NUMBER: 31,815  
CC REFERENCE/DOCKET NUMBER: P31 8717  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (619) 535-9001  
CC TELEFAX: (619) 535-8949  
CC INFORMATION FOR SEQ ID NO: 27:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 788 amino acids  
CC TYPE: AMINO ACID  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 788 AA; 85975 MW; 3136866 CN;  
Query Match 34.9%; Score 96; DB 2; Length 788;  
Best Local Similarity 42.9%; Pred. No. 6.70e+00;  
Matches 12; Conservative 11; Mismatches 4; Indels 1; Gaps 1;  
Db 706 PNIPMIMLVSLATLL-IGVLLCIWKL 732  
QY 227 PKLSKFISSLAILLMVSLLLSLWKL 254

CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 505  
CC TYPE: amino acid  
CC STRANDEDNESS: linear  
CC TOPOLOGY: linear  
CC SEQUENCE 505 AA; 56774 MW; 1307601 CN;  
Query Match 34.9%; Score 96; DB 1; Length 505;  
Best Local Similarity 44.0%; Pred. No. 6.70e+00;  
Matches 11; Conservative 10; Mismatches 4; Indels 0; Gaps 0;  
Db 122 LILGPVLALLVLVALGTGLRWRR 146  
QY 232 FILISSLAILLMVSLLLSLWKLWR 256

Query Match 34.9%; Score 96; DB 1; Length 505;  
Best Local Similarity 44.0%; Pred. No. 6.70e+00;  
Matches 11; Conservative 10; Mismatches 4; Indels 0; Gaps 0;





Db 93 IILAVIAGILMLALVFLWK 114  
:|:|:| | | :|:|:|:|:| | | |  
QY 232 FILISSLAIIIMVSLLLSLWK 253

RESULT 14  
ID US-08-241-387-7 STANDARD; PRT: 149 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX

Sequence 7, Application US/08241387  
Sequence 7, Application US/08241387  
Patent No. 5589570  
GENERAL INFORMATION:  
APPLICANT: Tamura, Richard N.  
APPLICANT: Quaranta, Vito  
TITLE OF INVENTION: INTEGRIN ALPHA SUBUNIT CYTOPLASMIC  
TITLE OF INVENTION: DOMAIN POLYPEPTIDES, ANTIBODIES AND METHODS  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute  
STREET: 10666 No. 5589570th Torrey Pines Road, TPC-8  
CITY: La Jolla  
STATE: CA  
COUNTRY: US  
ZIP: 92037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/241,387  
FILING DATE: 10-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 07/695,564  
FILING DATE: 03-MAY-1004  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI241.0D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-554-2937  
TELEFAX: 619-554-6312  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 149 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: Protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..149  
/note= "the 149 amino acid sequence  
OTHER INFORMATION:  
OTHER INFORMATION: predicted from the product which results from  
OTHER INFORMATION: amplification of the mouse ALPHA 6A cDNA with  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..120  
/note= "SEQ ID NO:7 is identical to  
OTHER INFORMATION:  
OTHER INFORMATION: SEQ ID NO:5 at amino acid positions 1 through 120;  
OTHER INFORMATION: the sequences diverge at amino acid 121."  
SEQUENCE 149 AA; 16829 MW; 116596 CN;  
Query Match 34.5%; Score 95; DB 1; Length 149;  
Best Local Similarity 50.0%; Pred. No. 7.75e+00;

Matches 11: Conservative 8; Mismatches 3; Indels 0; Gaps 0;  
Db 93 IILAVIAGILMLALVFLWK 114  
:|:|:| | | :|:|:|:|:| | | |  
QY 232 FILISSLAIIIMVSLLLSLWK 253  
RESULT 15  
ID 5196511-2 STANDARD; PRT: 1125 AA.  
XX  
AC xxxxxx  
XX  
DT 01-JAN-1900  
XX  
DE Patent No. 5196511.  
CC  
CC Patent No. 5196511  
CC APPLICANT: PLOW, EDWARD F.;D'SOUZA, STANLEY E.  
CC GINSBERG, MARK H.  
CC TITLE OF INVENTION: PEPTIDES AND ANTIBODIES THAT INHIBIT  
CC INTEGRIN-LIGAND BINDING  
CC NUMBER OF SEQUENCES: 31  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/07/444,777  
CC FILING DATE: 01-DEC-1989  
CC SEQ ID NO:2  
CC LENGTH: 1039  
CC SEQUENCE 1125 AA; 122832 MW; 7121651 CN;

Query Match 34.5%; Score 95; DB 4; Length 1039;  
Best Local Similarity 36.4%; Pred. No. 7.75e+00;  
Matches 8; Conservative 12; Mismatches 2; Indels 0; Gaps 0;  
Db 1000 VLVGVLGILLTILVIMWKV 1021  
:|:|:| | | :|:|:|:|:| | | |  
QY 233 ILISSLAIIIMVSLLLSLWK 254

Search completed: Wed May 10 12:02:34 2000  
Job time : 54 secs.





\*\*\*\*\*  
M P S R L H  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MParch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 12:00:47 2000; MasPar time 5.58 Seconds  
296.041 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-376-430-2  
Description: (226-260) from US09376430A.pep (7 of 25)  
Perfect Score: 275  
Sequence: 1 RPKLSKFLISSAILLMVSLLLSLWLWRVKKF 35

Scoring table: PAM 150  
Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pif62  
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 36.500; Variance 111.776; scale 0.327

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
1	106	38.5	108	2	S40149 integrin alpha-7C chain	3.70e-01
2	106	38.5	1106	2	S38783 integrin alpha chain	3.70e-01
3	105	38.2	1135	2	I93186 alpha-7 integrin mo	4.65e-01
4	104	37.8	1151	2	A45226 integrin alpha-1 subu	5.83e-01
5	104	37.8	1180	2	A38834 integrin alpha-1 chain	5.83e-01
6	102	37.1	1386	1	Q8E850 latent membrane prote	9.13e-01
7	102	37.1	1062	2	JC5951 integrin alpha 7 chain	9.13e-01
8	102	37.1	1137	2	JC5950 integrin alpha 7 chain	9.13e-01
9	101	36.7	381	2	S24611 latent membrane prote	1.14e+00
10	101	36.7	404	1	LABECA latent membrane prote	1.14e+00
11	101	36.7	418	2	D70038 maltodextrin transpor	1.14e+00
12	100	36.4	169	2	S76289 hypothetical protein	1.43e+00
13	100	36.4	183	2	I59442 receptor tyrosine kin	1.43e+00
14	100	36.4	220	2	A49508 receptor tyrosine kin	1.43e+00
15	100	36.4	876	2	A49508 protein-tyrosine kina	1.43e+00
16	100	36.4	910	2	A53137 tyrosine kinase recep	1.43e+00
17	100	36.4	913	2	A48280 receptor tyrosine kin	1.43e+00
18	100	36.4	1037	2	A60163 glycoprotein IIB - ra	1.43e+00
19	99	36.0	386	1	UABEB7 latent membrane prote	1.78e+00
20	99	36.0	471	2	A56839 integrin alpha-7 chain	1.78e+00
21	99	36.0	1073	2	B36429 integrin alpha-6 chain	1.78e+00
22	99	36.0	1091	2	A41543 integrin alpha-6 chain	1.78e+00
23	98	35.6	148	2	S40148 integrin alpha-7A chain	2.22e+00

24	97	35.3	129	2	I61187 alpha-7A integrin - m	2.76e+00
25	96	34.9	656	2	JC2005 integrin beta-5 chain	3.44e+00
26	96	34.9	788	2	A37057 integrin beta-6 chain	3.44e+00
27	96	34.9	799	2	A38308 integrin beta-5 chain	3.44e+00
28	95	34.5	141	2	A40463 integrin alpha-6 chain	4.27e+00
29	95	34.5	149	2	B40463 integrin alpha-6 chain	4.27e+00
30	95	34.5	604	2	I36917 glycoprotein IIB - ye	4.27e+00
31	95	34.5	1039	2	A34269 integrin alpha-2b cha	4.27e+00
32	95	34.5	2410	1	JQ1948 genome polyprotein 1	4.27e+00
33	94	34.2	1286	2	A33468 probable membrane pro	5.30e+00
34	94	34.2	1299	2	S06119 membrane protein patc	5.30e+00
35	93	33.8	549	1	H65214 hypothetical 60.5 kD	6.57e+00
36	93	33.8	858	1	JC2481 S-receptor kinase (EC	6.57e+00
37	92	33.5	127	2	S01397 H+-transporting ATP s	8.13e+00
38	92	33.5	133	2	PH0210 hypothetical protein	8.13e+00
39	92	33.5	614	2	A69845 Na+/H+ antiporter hom	8.13e+00
40	91	33.1	212	2	S13293 KDEL receptor - human	1.01e+01
41	91	33.1	212	2	A44394 KDEL receptor - bovin	1.01e+01
42	91	33.1	498	2	S11246 LAG-3 protein precurs	1.01e+01
43	91	33.1	503	2	A49451 activin/TGF-beta-like	1.01e+01
44	91	33.1	1327	2	T09402 immunoglobulin-like p	1.01e+01
45	90	32.7	547	2	F72042 integral membrane pro	1.24e+01

ALIGNMENTS

RESULT 1 S40149 #type complete  
ENTRY integrin alpha-7C chain - rat  
TITLE #formal name Rattus norvegicus #common\_name Norway rat  
ORGANISM 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change  
DATE 29-Sep-1999  
ACCESSIONS S40149  
REFERENCE S40147  
#authors Song, W.K.; Wang, W.; Sato, H.; Bielser, D.; Kaufman, S.  
#submission Submitted to the EMBL Data Library, July 1993  
#description Expression of alpha 7 integrin cytoplasmic domains during skeletal muscle development: alternate forms, conformational change, and homology with serine/threonine kinases and tyrosine phosphatases.  
#accession S40149  
##status preliminary  
##molecule\_type mRNA  
##residues 1-108 ##label SON  
##cross-references EMBL:X74294; NID:g437912; PIDN:CAA52347.1;  
PID:g437913

CLASSIFICATION #superfamily integrin alpha-2b chain  
SUMMARY #length 108 #molecular-weight 12016 #checksum 463

Query Match 38.5%; Score 106; DB 2; Length 108;  
Best Local Similarity 59.1%; Pred. No. 3.70e-01;  
Matches 13; Conservative 6; Mismatches 3; Indels 0; Gaps 0;  
Db 71 ILAVLAVGLLVALLVLLWLK 92  
QY 233 ILISSLAILLVSLLSLWLK 254  
|||||:||||:|||||

RESULT 2 S38783 #type fragment  
ENTRY integrin alpha chain - rat (fragment)  
TITLE #formal name Rattus norvegicus #common\_name Norway rat  
ORGANISM 22-Jan-1994 #sequence\_revision 14-Jul-1995 #text\_change  
DATE 29-Sep-1999

ACCESSIONS S38783; S23600  
REFERENCE S38783  
#authors Kaufman, S.J.  
#submission submitted to the EMBL Data Library, March 1992  
#accession S38783  
##status preliminary  
##molecule\_type mRNA  
##residues 1-1106 ##label KAU  
##cross-references EMBL:X65036; NID:g56392; PIDN:CAA46170.1; PID:g56393

```
##residues      1-1151 ##label BRI
##experimental_source hepatocblastoma cell line HepG2
##note          sequence extracted from NCBI backbone (NCBIP:124326)
CLASSIFICATION #superfamily unassigned collagens; von Willebrand factor type
                A repeat homology
FEATURE
  142-317      #domain von Willebrand factor type A repeat homology
              #label VWAI
SUMMARY        #length 1151 #checksum 3253

Query Match    37.8%; Score 104; DB 2; Length 1151;
Best Local Similarity 54.5%; Pred. No. 5.83e-01;
Matches 12; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

Db   1117 ILLSAFAGLLLLLMLLIILWKI 1138
QY   233 ILISSAAILLMVSLLSLWKL 254
       ||:::||::||::||::||::|
RESULT      5
ENTRY      A35854 #type complete
TITLE      Integrin alpha-1 chain precursor - rat
ORGANISM   #format_name Rattus norvegicus #common_name Norway rat
DATE       23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change
           20-Sep-1999
ACCESSIONS A35854; S11243
REFERENCE   A35854
#authors    Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton,
            A.; Esch, F.; Carbonetto, S.; Reichardt, L.F.
#journal     J. Cell Biol. (1990) 111:709-720
#title       Molecular cloning of the rat integrin alpha-1-subunit: a
            receptor for laminin and collagen.
#cross-references MUID:90338125
#accession    A35854
##status      preliminary
##molecule_type mRNA
##residues    1-1180 #label IGN
##cross-references GB:X52140; NID:g56493; PIDN:CAA36384.1; PID:g56494
CLASSIFICATION #superfamily unassigned collagens; von Willebrand factor type
                A repeat homology
KEYWORDS      cell adhesion; cytoskeleton; transmembrane protein
FEATURE
  170-345      #domain von Willebrand factor type A repeat homology
              #label VWAZ
SUMMARY        #length 1180 #molecular-weight 130808 #checksum 6967

Query Match    37.8%; Score 104; DB 2; Length 1180;
Best Local Similarity 54.5%; Pred. No. 5.83e-01;
Matches 12; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

Db   1146 ILLSAFAGLLLLLMLLIILWKI 1167
QY   233 ILISSAAILLMVSLLSLWKL 254
       ||:::||::||::||::||::|
RESULT      6
ENTRY      QBSE0 #type complete
TITLE      Latent membrane protein - human herpesvirus 4 (strain B95-8)
ORGANISM   #format_name Human herpesvirus 4, Epstein-Barr virus
DATE       25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change
           16-Feb-1997
ACCESSIONS    D43045; A03794
REFERENCE      A93065
#authors       Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
#journal       Mol. Biol. Med. (1983) 1:21-45
#title         Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8
            virus.
#cross-references MUID:85035713
#accession     D43045
##molecule_type DNA
##residues    1-386 #label BAN
REFERENCE      A03794
#authors       Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.
```

Farrell, P.J.; Gibson, T.J.; Hatfull, G.; Hudson, G.S.; Satchwell, S.C.; Seguin, C.; Tuffnell, P.S.; Barrall, B.G. Nature (1984) 310:207-211  
 #journal #title  
 #cross-references MUID:84270667  
 #contents annotation; protein coding region  
 #CLASSIFICATION #superfamily Epstein-Barr virus latent membrane protein transmembrane protein  
 #KEYWORDS #length 386 #molecular-weight 41982 #checksum 6693

Query Match 37.1%; Score 102; DB 1; Length 386;  
 Best Local Similarity 52.4%; Pred. No. 9.13e-01;  
 Matches 11; Conservative 8; Mismatches 1; Indels 1; Gaps 1;  
 Db 80 LGALCILLMILLIALLNL 100  
 QY 235 ISSLAILL-MVSLLSLWKL 254

RESULT 7  
 ENTRY JC5951 #type complete  
 TITLE Integrin alpha 7 chain variant - human  
 ORGANISM Homo sapiens #common\_name man  
 DATE 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 17-Mar-1999  
 ACCESSIONS JC5951  
 REFERENCE Leung, E.; Lim, S.P.; Berg, R.; Yang, Y.; Ni, J.; Wang, S.; Krissansen, G.W.  
 #journal Biochem. Biophys. Res. Commun. (1998) 243:317-325  
 #title A novel extracellular domain variant of the human integrin alpha 7 subunit generated by alternative intron splicing.  
 #cross-references MUID:98139911  
 #accession JC5951  
 #status preliminary  
 #molecule\_type mRNA  
 #residues 1-1062 #label LEU  
 #cross-references GB:AF032108  
 #CLASSIFICATION #superfamily integrin alpha-2b chain  
 #SUMMARY #length 1062 #molecular-weight 116404 #checksum 2496

Query Match 37.1%; Score 102; DB 2; Length 1062;  
 Best Local Similarity 54.5%; Pred. No. 9.13e-01;  
 Matches 12; Conservative 7; Mismatches 3; Indels 0; Gaps 0;  
 Db 966 ILLAVLAGLLVLLVLLWKM 987  
 QY 233 ILLSAILLMVSLLSLWKL 254

RESULT 8  
 ENTRY JC5950 #type complete  
 TITLE Integrin alpha 7 chain - human  
 ORGANISM Homo sapiens #common\_name man  
 DATE 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 29-Sep-1999  
 ACCESSIONS JC5950  
 REFERENCE Leung, E.; Lim, S.P.; Berg, R.; Yang, Y.; Ni, J.; Wang, S.; Krissansen, G.W.  
 #journal Biochem. Biophys. Res. Commun. (1998) 243:317-325  
 #title A novel extracellular domain variant of the human integrin alpha 7 subunit generated by alternative intron splicing.  
 #cross-references MUID:98139911  
 #accession JC5950  
 #status preliminary  
 #molecule\_type mRNA  
 #residues 1-1137 #label LEU  
 #cross-references GB:AF032109; NID:g2897115; PIDN:AC39708.1; PID:g2897116

CLASSIFICATION #superfamily integrin alpha-2b chain  
 #SUMMARY #length 1062 #molecular-weight 116404 #checksum 2496

Query Match 37.1%; Score 102; DB 2; Length 1062;  
 Best Local Similarity 54.5%; Pred. No. 9.13e-01;  
 Matches 12; Conservative 7; Mismatches 3; Indels 0; Gaps 0;  
 Db 966 ILLAVLAGLLVLLVLLWKM 987  
 QY 233 ILLSAILLMVSLLSLWKL 254

RESULT 8  
 ENTRY JC5950 #type complete  
 TITLE Integrin alpha 7 chain - human  
 ORGANISM Homo sapiens #common\_name man  
 DATE 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 29-Sep-1999  
 ACCESSIONS JC5950  
 REFERENCE Leung, E.; Lim, S.P.; Berg, R.; Yang, Y.; Ni, J.; Wang, S.; Krissansen, G.W.  
 #journal Biochem. Biophys. Res. Commun. (1998) 243:317-325  
 #title A novel extracellular domain variant of the human integrin alpha 7 subunit generated by alternative intron splicing.  
 #cross-references MUID:98139911  
 #accession JC5950  
 #status preliminary  
 #molecule\_type mRNA  
 #residues 1-1137 #label LEU  
 #cross-references GB:AF032109; NID:g2897115; PIDN:AC39708.1; PID:g2897116

CLASSIFICATION #superfamily integrin alpha-2b chain  
 #SUMMARY #length 1137 #molecular-weight 124287 #checksum 6281

Query Match 37.1%; Score 102; DB 2; Length 1137;  
 Best Local Similarity 54.5%; Pred. No. 9.13e-01;  
 Matches 12; Conservative 7; Mismatches 3; Indels 0; Gaps 0;  
 Db 1041 ILLAVLAGLLVLLVLLWKM 1062

QY 233 ILLSAILLMVSLLSLWKL 254

RESULT 9  
 ENTRY S24611 #type complete  
 TITLE Latent membrane protein human herpesvirus 4  
 ORGANISM Homo sapiens #formal\_name human herpesvirus 4 Epstein-Barr virus  
 DATE 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

ACCESSIONS S24611  
 REFERENCE Chang, Y.S.  
 #authors submitted to the EMBL Data Library, June 1992  
 #submission S24611  
 #accession S24611  
 #status preliminary  
 #molecule\_type DNA  
 #residues 1-381 #label CHA  
 #cross-references EMBL:X66863; NID:g59181; PIDN:CAA47332.1; PID:g59182

GENETICS 90/1; 119/1  
 #introns  
 CLASSIFICATION #superfamily Epstein-Barr virus latent membrane protein  
 #SUMMARY #length 381 #molecular-weight 41372 #checksum 3544

Query Match 36.7%; Score 101; DB 2; Length 381;  
 Best Local Similarity 47.6%; Pred. No. 1.14e+00;  
 Matches 10; Conservative 10; Mismatches 0; Indels 1; Gaps 1;  
 Db 80 LGGGLLLMILLIALLNL 100  
 QY 235 ISSLAILL-MVSLLSLWKL 254

RESULT 10  
 ENTRY LABECA #type complete  
 TITLE Latent membrane protein LMP1 - human herpesvirus 4 (strain CAO)  
 ORGANISM Homo sapiens #formal\_name human herpesvirus 4, Epstein-Barr virus  
 DATE 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 16-Jul-1999  
 ACCESSIONS J01434; G00065; S21660  
 REFERENCE J01434  
 #authors Hu, L.F.; Zabarovsky, E.R.; Chen, F.; Cao, S.L.; Ernberg, I.; Klein, G.; Winberg, G.

#journal J. Gen. Virol. (1991) 72:2399-2409  
 #title Isolation and sequencing of the Epstein-Barr virus BNLF-1 gene (LMP1) from a Chinese nasopharyngeal carcinoma.  
 #cross-references MUID:92013956  
 #accession J01434  
 #molecule\_type DNA  
 #residues 1-404 #label HUI  
 #cross-references EMBL:X58140; NID:g22937; PIDN:CAA41148.1; PID:g22938  
 #note The authors translated the codon AAA for residue 358 as Ala

#note warning: the organism was shown incorrectly as Homo sapiens in GenBank accession X58140, release 100.0  
 Unlike Epstein-Barr nuclear antigen 1 (EBNA-1) (see PIR:S42440), which is expressed in all EBV-infected cells, this protein appears associated with transformation and is found in a majority of nasopharyngeal carcinoma tumor biopsies.

GENETICS LMP1; BNLF-1  
 #gene 90/1; 119/1  
 #introns  
 CLASSIFICATION #superfamily Epstein-Barr virus latent membrane protein  
 #KEYWORDS oncogene; tandem repeat; transmembrane protein  
 #FEATURE #domain transmembrane #status predicted #label TM1  
 25-45

```

GENETICS
#gene      yvfL
CLASSIFICATION #superfamily inner membrane protein malf
SUMMARY     #length 418 #molecular-weight 46578 #checksum 6122

Query Match          36.7%;   Score 101;   DB 2;   Length 418;
Best Local Similarity 35.5%;   Pred. NO. 1.14e+00;
Matches 11; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

Db    384 SOYSLAALTLILSVFVSIASIALWQFRKTF 414
      | : : |:|:| | : :|:| : :| :|
QY    230 SKFFILISLAIIAMVSLLELLSLWKLRVKKF 260

```

RESULT	12
ENTRY	S76289
TITLE	#type complete hypothetical protein sll0751 - <i>Synechocystis</i> sp. (strain PCC 6803)
ORGANISM	#formal_name <i>Synechocystis</i> sp.
#variety	PCC 6803
DATE	25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 21-Aug-1998
ACCESSIONS	S76289
REFERENCE	S74322
#authors	Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, K.

Nakamura, Y.; Miyajima, N.; Hiroseawa, M.; Sugizawa, M.;  
 Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.;  
 Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimpo,  
 S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.;  
 Yasuda, M.; Tabata, S.  
 DNA Res. (1996) 3:109-136  
 Sequence analysis of the genome of the unicellular  
 cyanobacterium *Synechocystis* sp. PCC6803. II. Sequence  
 determination of the entire genome and assignment of  
 potential protein-coding regions.  
 #cross-references MURD:97061201  
 #accession S76289  
 #status  
 #molecule\_type DNA  
 #residues 1-169 #label KAN  
 #cross-references EMBL:D64000; GB:AB001339; NID:g1001484; PID:g1010792;  
 PID:g1001514  
 the nucleotide sequence was submitted to the EMBL Data  
 Library, June 1996  
 ##note

```
#sequence_revision 02-Aug-1996 #text_change
#cross-references PMID:95320273
#accession I59442
#status preliminary; translated from GB/EMBL/DDBJ
#molecule_type mRNA
#residues_1-183 #label RES
#cross-references GB:S77596; NID:g957334; PIDN:AAB34729.1 PID:g957335
#authors Sakuma, S.; Saya, H.; Ijichi, A.; Tofilon, P.J.
#journal Radiat. Res. (1995) 143:1-7
#title Radiation induction of the receptor tyrosine kinase gene
          Ptk-3 in normal rat astrocytes.
```

```

FEATURE
571-185      #domain discoidin I amino-terminal homology #label DNI\
571-875      #domain protein kinase homology #label KIN\
579-587      #region protein kinase ATP-binding motif
SUMMARY      #length 876 #molecular-weight 971173 #checksum 5186

Query Match      36.4%; Score 100; DB 2; Length 876;
Best Local Similarity 51.9%; Fied. NO. 1.43e+00;
Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;

Ddb 418 ILIGCLVAIIELLLLIIALLMLRWLR 444
      ||| | |||::|: | ||| ||
Qy 233 ILISSL-AILLMVSLLL-LSLWKL-WR 256
      †
Search completed: Wed May 10 12:00:56 2000
Job time : 9 secs.

```

Search completed: Wed May 10 12:00:56 2000  
Job time : 9 secs.

GENETICS

#gene  
#CLASSIFICATION  
#length 183 #checksum 5641

SUMMARY

Query Match 36.4%; Score 100; DB 2; Length 183;  
Best Local Similarity 51.9%; Pred. No. 1.43e+00;  
Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;

DBD 15 ILIGCLVAIIIIIIALMLRWLR 41  
QY 233 ILISSL-AIIIMVSLLL-LSLWKL-WR 256

RESULT 14

ENTRY #type fragment  
#receptor tyrosine kinase - rat (fragment)  
#formal\_name Rattus sp. #common\_name rat  
#organism  
#date 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 24-Sep-1999

ACCESSIONS R80329  
REFERENCE R59442  
#authors Sakuma, S.; Sava, H.; Ijichi, A.; Tofilon, P.J.  
#journal Radiat. Res. (1995) 143:1-7  
#title Radiation induction of the receptor tyrosine kinase gene  
#ptk-3 in normal rat astrocytes.  
#cross-references MUID:95320273

#accession R80329  
#status Preliminary; translated from GB/EMBL/DDBJ  
#molecule\_type mRNA  
#residues 1-220 #label RES  
#cross-references GB:S77585; NID:g957332; PIDN:AAB34728.1; PID:g957333

GENETICS

#gene  
#CLASSIFICATION  
#length 220 #checksum 1895

SUMMARY

Query Match 36.4%; Score 100; DB 2; Length 220;  
Best Local Similarity 51.9%; Pred. No. 1.43e+00;  
Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;

DB 15 ILIGCLVAIIIIIIALMLRWLR 41  
QY 233 ILISSL-AIIIMVSLLL-LSLWKL-WR 256

RESULT 15

ENTRY #type complete  
#protein-tyrosine kinase (EC 2.7.1.112) trke precursor - human  
#formal\_name Homo sapiens #common\_name man  
#organism  
#date 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 24-Sep-1999

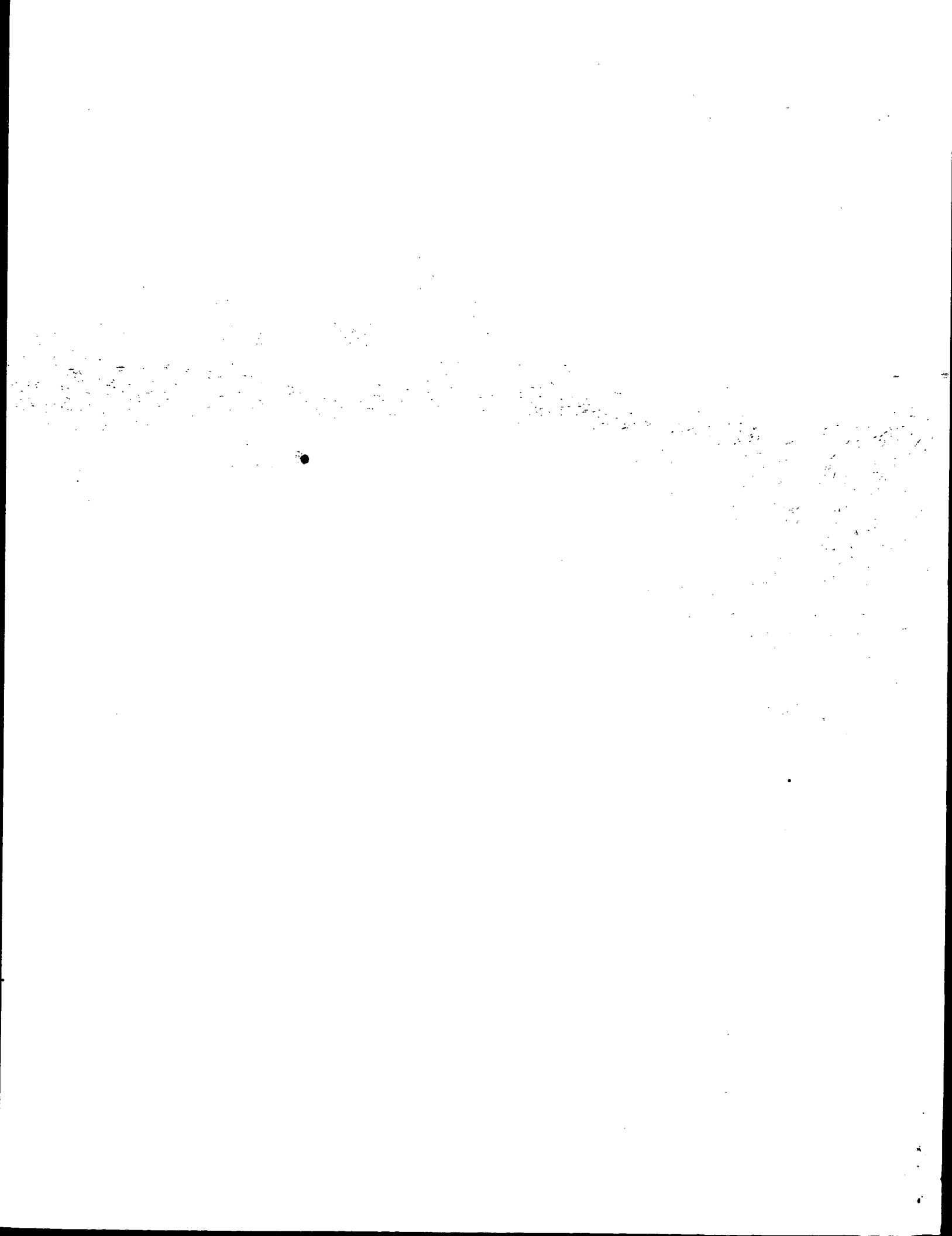
ACCESSIONS A49508; I38358; S37402  
REFERENCE A49508  
#authors Di Marco, E.; Cutuli, N.; Guerra, L.; Cancedda, R.; De Luca, M.

#journal J. Biol. Chem. (1993) 268:24290-24295  
#title Molecular cloning of trke, a novel trk-related putative tyrosine kinase receptor isolated from normal human keratinocytes and widely expressed by normal human tissues

#cross-references MUID:9403285  
#accession A49508  
#status Preliminary  
#molecule\_type mRNA  
#residues 1-876 #label DIA  
#cross-references EMBL:X74979; NID:g400462; PIDN:CAA52915.1; PID:g400463

CLASSIFICATION #superfamily unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino-terminal homology; protein kinase homology  
ATP; phosphotransferase

KEYWORDS



\*\*\*\*\*  
M P S R L H (TM)  
\*\*\*\*\*  
Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd  
Mpsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 11:54:33 2000; MasPar time 92.69 seconds  
Tabular output not generated. 11.500 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (226-260) from US09376430A.ppt (7 of 25)  
Perfect score: 275  
Sequence: 1 KPGLSKFILSSLAALLMVSLLLSLKLNVRKF 35  
Scoring table: PAM 150  
Gap 11  
Searched: 83857 seqs, 30454973 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: swiss-prot38  
1:swissprot  
Statistics: Mean 37.179; Variance 101.980; scale 0.365  
pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Pred. No.	
Result No.	Score	Query Match	Description	ID	
1	104	37.8	INTEGRIN ALPHA-1 (LAMI	ITAI_HUMAN	1.42e-01
2	104	37.8	INTEGRIN ALPHA-1 (LAMI	ITAI_HUMAN	1.42e-01
3	102	37.1	LATENT MEMBRANE PROTEI	LMP1_EBVC	2.32e-01
4	101	36.7	LATENT MEMBRANE PROTEI	LMP1_EBVC	2.32e-01
5	100	36.4	EPITHELIAL DISCOIDIN D	EDD1_MOUSE	3.79e-01
6	100	36.4	EPITHELIAL DISCOIDIN D	EDD1_MOUSE	3.79e-01
7	100	36.4	EPITHELIAL DISCOIDIN D	EDD1_MOUSE	3.79e-01
8	99	36.0	LATENT MEMBRANE PROTEI	LMP1_EBVR	4.83e-01
9	99	36.0	LATENT MEMBRANE PROTEI	LMP1_EBVR	4.83e-01
10	96	34.9	SERINE/THREONINE-PROTE	KIR3_RAT	9.95e-01
11	96	34.9	SERINE/THREONINE-PROTE	KIR3_RAT	9.95e-01
12	96	34.9	INTEGRIN BETA-5 (FRAGM	ITB5_PAPCY	9.95e-01
13	96	34.9	INTEGRIN BETA-5 (FRAGM	ITB5_PAPCY	9.95e-01
14	95	34.5	PLATELET MEMBRANE GLYC	ITAB_HUMAN	1.26e+00
15	95	34.5	PLATELET MEMBRANE GLYC	ITAB_HUMAN	1.26e+00
16	95	34.5	PLATELET MEMBRANE GLYC	ITAB_HUMAN	1.26e+00
17	94	34.2	GENOME POLYPROTEIN 1	POL1_BAYM	1.26e+00
18	94	34.2	GENOME POLYPROTEIN 1	POL1_BAYM	1.26e+00
19	93	33.8	HEDGEHOG RECEPTOR (PAT	YICE_ECOLI	2.03e+00
20	92	33.5	ATP SYNTHASE PROTEIN 1	ATPZ_EACB3	2.56e+00
21	91	33.1	ER LUMEN PROTEIN RETAI	ERL1_BOVIN	3.24e+00
22	91	33.1	ER LUMEN PROTEIN RETAI	ERL1_BOVIN	3.24e+00
23	91	33.1	SERINE/THREONINE-PROTE	KIR3_HUMAN	3.24e+00

ALIGNMENTS				PRT; 1151 AA.	
RESULT ID	ITAI_HUMAN	STANDARD;			
AC	P56199;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	INTEGRIN ALPHA-1 (LAMININ AND COLLAGEN RECEPTOR) (VLA-1) (CD49A).				
GN	ITGAL				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
OC	Eutheria; Primates; Catarrhini; Hominidae; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	MEDLINE; 93155124.				
RA	Briesewitz R., Epstein M.R., Marcantonio E.E.;				
RT	"Expression of native and truncated forms of the human integrin alpha				
RT	1 subunit."				
RL	J. Biol. Chem. 268:2989-2996(1993).				
CC	-1- FUNCTION: ASSOCIATED WITH THE BETA SUBUNIT, THIS PROTEIN FORMS				
CC	A LAMININ AND COLLAGEN RECEPTOR.				
CC	-1- SUBUNIT: DIMER OF AN ALPHA AND BETA CHAINS. THE BETA CHAIN THAT				
CC	ASSOCIATES WITH ALPHA-1 IS BETA-1.				
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.				
CC	-1- DOMAIN: HAS A 206 AA INSERT (CALLED THE I-DOMAIN), THE SUBUNITS				
CC	WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.				
CC	-1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY. I-DOMAIN				
CC	SUBFAMILY.				
CC	-1- DATABASE: NAME=PROW; NOTE=CD guide CD49a entry;				
CC	WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49a.htm".				
DR	HSP; P11213; 1A8X.				
DR	MM; 192368;				
DR	PRIN5; PR00453; VWFADOMAIN.				
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.				
DR	PRAM; PF00357; vwa; 1.				
DR	PRAM; PF00357; Integrin_A; 1.				
KW	Integrin; Cell adhesion; Glycoprotein; Transmembrane;				
KW	Extracellular matrix; Cytoskeleton.				
FT	DOMAIN 1113				
FT	TRANSMEM 1114..1136				
FT	DOMAIN 1137..1151				
FT	DOMAIN 1147..360				
FT	DISULFID 54				
FT	DISULFID 660				
FT	DISULFID 675				
FT	DISULFID 728				
FT	DISULFID 780				
FT	DISULFID 850				
FT	DISULFID 1002				
FT	DISULFID 1034				

24	91	33.1	525	1	LAC3_HUMAN	LYMPHOCYTE ACTIVATION	3.24e+00
25	90	32.7	400	1	LEUK_HUMAN	LEUKOSIALIN PRECURSOR	4.08e+00
26	89	32.4	779	1	YHUN_ECOLI	CYTOCHROME C OXIDASE P	5.14e+00
27	88	32.0	261	1	COX3_LATCH	CYTOCHROME C OXIDASE P	6.46e+00
28	88	32.0	496	1	C4D2_DROME	CYTOCHROME P450 4D2 (E	6.46e+00
29	88	32.0	860	1	LDLR_HUMAN	LOW-DENSITY LIPOPROTEIN	6.46e+00
30	88	32.0	864	1	LDLR_MOUSE	LOW-DENSITY LIPOPROTEIN	6.46e+00
31	87	31.6	405	1	CYB_RHOU	CYTOCHROME B.	8.11e+00
32	86	31.6	875	1	AMD2_XENLA	PEPTIDYL-GLYCINE ALPHA	8.11e+00
33	86	31.3	351	1	AOX_ASPNG	ALTERNATIVE OXIDASE PR	1.02e+01
34	86	31.3	383	1	EDG1_RAT	PROBABLE G PROTEIN-COU	1.02e+01
35	86	31.3	908	1	PTGA_MYCGE	PTS SYSTEM, GLUCOSE-SP	1.02e+01
36	86	31.3	1207	1	EGF_HUMAN	PRO-EPIDERMAL GROWTH F	1.02e+01
37	85	30.9	111	1	USPB_ECOLI	UNIVERSAL STRESS PROTE	1.27e+01
38	85	30.9	120	1	YCV2_YEAST	HYPOTHETICAL 13.8 KD P	1.27e+01
39	85	30.9	161	1	GSPM_KLEPN	GENERAL SECRETION PATH	1.27e+01
40	85	30.9	490	1	CPB3_MOUSE	CYTOCHROME P450 2C38	1.27e+01
41	85	30.9	503	1	CPV1_BOVIN	CYTOCHROME P450 19 (AR	1.27e+01
42	85	30.9	511	1	YNUC_ECOLI	HYPOTHETICAL 56.2 KD P	1.27e+01
43	85	30.9	773	1	PAC2_PSESS	PENICILLIN ACYLASE II	1.27e+01
44	85	30.9	814	1	CADF_HUMAN	MUSCLE-CADHERIN PRECUR	1.27e+01
45	85	30.9	1166	1	AMRK_XIPMA	MELANOMA RECEPTOR PRO	1.27e+01





```

RX MEDLINE; 84270667.
RA Beer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hafell G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Turfitt P.S., Barrell B.G.;
RA "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RA Nature 310:207-211(1984).
RL [2]
RN SEQUENCE FROM N.A.
RP MEDLINE; 85108145.
RX Hudson G.S., Farrell P.J., Barrell B.G.;
RA "Two related but differentially expressed potential membrane proteins
RT encoded by the EcoRI Dhet region of Epstein-Barr virus B95-8.";
RL J. Virol. 53:528-535(1985).
RN [3]
RP IDENTIFICATION OF PROTEIN.
RX MEDLINE; 87112999.
RA Baichwal V.R., Suddgen B.;
RT "Posttranslational processing of an Epstein-Barr virus-encoded
RT membrane protein expressed in cells transformed by Epstein-Barr
RT virus.";
RL J. Virol. 61:866-875(1987).
RN [4]
RP TRANSFORMING PROPERTIES.
RX MEDLINE; 88233636.
RA Baichwal V.R., Suddgen B.;
RT "Transformation of Balb 3T3 cells by the BNLF-1 gene of Epstein-Barr
RT virus.";
RL Oncogene 2:461-467(1988).
RN [5]
RP PROCESSING.
RX MEDLINE; 90112654.
RA Moorthy R., Thorley-Lawson D.A.;
RT "Processing of the Epstein-Barr virus-encoded latent membrane protein
RT p63/LMP.";
RL J. Virol. 64:829-837(1990).
RN [5]
RP FUNCTION: THE LATENT MEMBRANE PROTEIN HAS TRANSFORMING ACTIVITY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V01555; NOT ANNOTATED_CDS.
CC EMBL; X01995; CAA26023.1;
CC PIR; A03794; Q0BES0.
CC Transmembrane; Phosphorylation; Transforming protein.
CC PEPTIDE 242 386
CC DOMAIN 1 24 CYTOPLASMIC.
CC TRANSMEM 25 44 POTENTIAL.
CC TRANSMEM 52 72 POTENTIAL.
CC TRANSMEM 77 97 POTENTIAL.
CC TRANSMEM 105 125 POTENTIAL.
CC TRANSMEM 139 159 POTENTIAL.
CC TRANSMEM 166 186 POTENTIAL.
CC TRANSMEM 187 386 POTENTIAL.
CC DOMAIN 187 386 CYTOPLASMIC.
CC SEQUENCE 386 AA; 41982 MW; 1E19446E857DB8A3 CRC64;
CC
CC Query Match 37.1%; Score 102; DB 1; Length 386;
CC Best Local Similarity 52.4%; Pred. No. 2.32e-01;
CC Matches 11; Conservative 8; Mismatches 1; Indels 1; Gaps 1;
CC
CC Db 80 LGALCILLMITLLIALWNL 100
CC QY 235 ISSLAILL-MVSLLSLWKL 254
CC
CC RESULT 4
CC ID LMP1_EBVC STANDARD; PRT; 404 AA.
CC AC P29362;
CC DT 01-DEC-1992 (Rel. 24, Created)

```

```

DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DT LATENT MEMBRANE PROTEIN 1 (LMP-1) (P63).
GN BNLF1.
OS Epstein-Barr virus (strain Cao) (Human herpesvirus 4).
Virus; dsDNA viruses, no RNA stage: Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92013956.
RA Hu L.F., Zabarovsky E.R., Chen F., Cao S.L., Ernberg I., Klein G.,
RA Winberg G.;
RT "Isolation and sequencing of the Epstein-Barr virus BNLF-1 gene
RT (LMP1) from a Chinese nasopharyngeal carcinoma.";
RL J. Gen. Virol. 72:2399-2409(1991).
RN [1]
RP FUNCTION: THE LATENT MEMBRANE PROTEIN HAS TRANSFORMING ACTIVITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X58140; CAA41148.1;
CC EMBL; D10059; BAA00948.1;
CC PIR; J01434; IABECA.
CC PIR; S21660; S21660.
CC Transmembrane; Phosphorylation; Transforming protein.
CC DOMAIN 1 24 CYTOPLASMIC.
CC TRANSMEM 25 44 POTENTIAL.
CC TRANSMEM 52 72 POTENTIAL.
CC TRANSMEM 77 97 POTENTIAL.
CC TRANSMEM 105 125 POTENTIAL.
CC TRANSMEM 139 159 POTENTIAL.
CC TRANSMEM 166 186 POTENTIAL.
CC TRANSMEM 187 404 CYTOPLASMIC.
CC SEQUENCE 404 AA; 43769 MW; D04536D3B65FF82E CRC64;
CC
CC Query Match 36.7%; Score 101; DB 1; Length 404;
CC Best Local Similarity 47.6%; Pred. No. 2.97e-01;
CC Matches 10; Conservative 10; Mismatches 0; Indels 1; Gaps 1;
CC
CC Db 80 LGGLGILLMITLLIALWNL 100
CC QY 235 ISSLAILL-MVSLLSLWKL 254
CC
CC RESULT 5
CC ID EDD1_RAT STANDARD; PRT; 910 AA.
CC AC Q63474.
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DE EPIHELIAL DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR (EC 2.7.1.112)
CC DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE
CC DE DDR) (DISCOIDIN RECEPTOR TYROSINE KINASE) (PROTEIN-TYROSINE KINASE
CC DE PTK-3).
CC GN EDD1 OR PTK3.
CC OS Rattus norvegicus (Rat).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
CC OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-SPRAGUE-DAWLEY; TISSUE=BRAIN;
RA MEDLINE; 94173920.
RA Sanchez M.P., Tapley P., Saini S.S., He B., Pulido D., Barbacid M.;
RT "Multiple tyrosine protein kinases in rat hippocampal neurons:
RT isolation of ptk-3 a receptor expressed in proliferative zones of
RT the developing brain";
RT Proc. Natl. Acad. Sci. U.S.A. 91:1819-1823(1994).
RN [1]
RP FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND

```



FT NP\_BIND 614 622 ATP (BY SIMILARITY).  
 FT BINDING 653 653 ATP (BY SIMILARITY).  
 FT ACT\_SITE 764 764 BY SIMILARITY.  
 FT DISULFID 32 186 BY SIMILARITY.  
 FT MOD\_RES 511 511 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 790 790 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 794 794 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 795 795 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 213 213 POTENTIAL.  
 FT CARBOHYD 262 262 POTENTIAL.  
 FT CARBOHYD 372 372 POTENTIAL.  
 FT CARBOHYD 392 392 POTENTIAL.  
 FT VARSPLIC 503 539 MISSING (IN ISOFORM CAK II).  
 SQ SEQUENCE 911 AA; 101160 MW; DBB7FE03DDDD79510 CRC64;

Query Match 36.4%; Score 100; DB 1; Length 911;  
 Best Local Similarity 51.9%; Pred. NO. 3.79e-01;  
 Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;

Db 416 ILIGLVAIIILLIIALLMLWRLHWR 442  
 QY 233 ILISSL-ALLMVSLLL-LSLWKL-WR 256

RESULT 7  
 ID EDDL\_HUMAN STANDARD; PRT; 913 AA.  
 AC Q08345; Q14196; Q16362;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE EPITHELIAL DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR (EC 2.7.1.112)  
 DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE  
 DE (DISCOIDIN RECEPTOR TYROSINE KINASE) (TRK E) (PROTEIN-TYROSINE  
 DE KINASE RTK 6)  
 GN EDDL OR CAK OR DDR1 OR TRKE OR RTK6.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RX MEDLINE; 93296201.  
 RA Johnson J.D., Edman J.C., Rutter W.J.;  
 RA "A receptor tyrosine kinase found in breast carcinoma cells has an  
 RT extracellular discoidin I-like domain.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 90:5677-5681(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RX MEDLINE; 93296201.  
 RA Johnson J.D., Edman J.C., Rutter W.J.;  
 RA "A receptor tyrosine kinase found in breast carcinoma cells has an  
 RT extracellular discoidin I-like domain.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 90:5677-5681(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE; 97131588.  
 RA Sakuma S., Tada M., Sava H., Sawamura Y., Shinohara Y., Abe H.;  
 RA "Receptor protein tyrosine kinase DDR is up-regulated by p53  
 RT protein.";  
 RT FEBS Lett. 398:165-169(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE; 96389017.  
 RA Playford M.P., Butler R.J., Wang X.C., Katso R.M., Cooke I.E.,  
 RA Ganesan T.S.;  
 RA "The genomic structure of discoidin receptor tyrosine kinase.";  
 RL Genome Res. 6:620-627(1996).  
 RN [5]  
 RP SEQUENCE FROM N.A. (SHORT FORM).  
 RC TISSUE=LUNG;  
 RX MEDLINE; 96204002.

RA Perez J.L., Jing S.O., Wong T.W.;  
 RT "Identification of two isoforms of the Cak receptor kinase that are  
 RL coexpressed in breast tumor cell lines.";  
 RN Oncogene 12:1469-1477(1996).  
 RP SEQUENCE FROM N.A. (SHORT FORM).  
 RC TISSUE=BRAIN, AND KERATINOCYTES;  
 RX MEDLINE; 94043265.  
 RA di Marco E., Cutuli N., Guerra L., Cancedda R., de Luca M.;  
 RT "Molecular cloning of trke, a novel trk-related putative tyrosine  
 RL kinase receptor isolated from normal human keratinocytes and widely  
 RN expressed by normal human tissues.";  
 RP J Biol. Chem. 268:24290-24295(1993).  
 RN [7]  
 RP SEQUENCE FROM N.A. (SHORT FORM).  
 RC TISSUE=OVARY;  
 RX MEDLINE; 95151639.  
 RA Laval S., Butler R., Shelling A.N., Hanby A.M., Poulson R.,  
 RA Ganesan T.S.;  
 RT "Isolation and characterization of an epithelial-specific receptor  
 RL tyrosine kinase from an ovarian cancer cell line.";  
 RN Cell Growth Differ. 5:1173-1183(1994).  
 CC -!- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND  
 CC RECOGNITION.  
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +  
 CC PROTEIN TYROSINE PHOSPHATE.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS; CAK I (SHOWN HERE) AND CAK II;  
 CC ARE PRODUCED BY ALTERNATIVE SPLICING. THEY ONLY DIFFER BY THE  
 CC ABSENCE OF A 37 RESIDUES SEGMENT.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN MOST ADULT TISSUES  
 CC AND IS HIGHEST IN THE BRAIN AND LUNG. ABUNDANT IN BREAST CARCINOMA  
 CC CELL LINES.  
 CC -!- DOMAIN: THE GLY/PRO-RICH DOMAINS MAY BE REQUIRED FOR AN UNUSUAL  
 CC GEOMETRY OF INTERACTION WITH LIGAND OR SUBSTRATES.  
 CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.  
 CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L11315; AAA02866.1;  
 DR EMBL; L20817; AAA18019.1;  
 DR EMBL; U48705; AAC50917.1;  
 DR EMBL; X98208; CAA66871.1;  
 DR EMBL; X99023; CAA66871.1; JOINED.  
 DR EMBL; X99024; CAA66871.1; JOINED.  
 DR EMBL; X99025; CAA66871.1; JOINED.  
 DR EMBL; X99026; CAA66871.1; JOINED.  
 DR EMBL; X99027; CAA66871.1; JOINED.  
 DR EMBL; X99028; CAA66871.1; JOINED.  
 DR EMBL; X99029; CAA66871.1; JOINED.  
 DR EMBL; X99030; CAA66871.1; JOINED.  
 DR EMBL; X99031; CAA66871.1; JOINED.  
 DR EMBL; X99032; CAA66871.1; JOINED.  
 DR EMBL; X99033; CAA66871.1; JOINED.  
 DR EMBL; X99034; CAA66871.1; JOINED.  
 DR EMBL; L57508; AAB05208.1;  
 DR EMBL; X74979; CAA52915.1;  
 DR EMBL; Z29093; CAA82335.1;  
 DR HSP; P11362; 1FGI.  
 DR MIM; 600452;  
 DR MIM; 600408;  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.

DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 DR PROSITE; PS01285; FA58C\_1; 1.  
 DR PROSITE; PS01286; FA58C\_2; 1.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PFAM; PF00754; F5\_F8\_type\_C; 1.  
 KW Transferase: Tyrosine-protein kinase; Glycoprotein; Signal;  
 KW Phosphorylation; Transmembrane; Receptor; ATP-binding;  
 KW Alternative splicing.  
 FT SIGNAL 1 18  
 FT CHAIN 19 913  
 FT DOMAIN 19 416  
 FT TRANSMEM 417 443  
 FT DOMAIN 444 913  
 FT DOMAIN 31 185  
 FT DOMAIN 377 415  
 FT DOMAIN 476 601  
 FT DOMAIN 610 905  
 FT NP\_BIND 616 624  
 FT BINDING 635 635  
 FT ACT\_SITE 766 766  
 FT DISULFID 31 185  
 FT MOD\_RES 513 513  
 FT MOD\_RES 792 792  
 FT MOD\_RES 796 796  
 FT MOD\_RES 797 797  
 FT CARBOHYD 211 211  
 FT CARBOHYD 260 260  
 FT CARBOHYD 371 371  
 FT CARBOHYD 394 394  
 FT VARSPLIC 506 542  
 FT CONFLICT 94 94  
 FT CONFLICT 833 833  
 FT CONFLICT 847 867  
 FT SEQUENCE 913 AA; 101127 MW; C96913EA906C481E CRC64;  
 Query Match 36.48; Score 100; DB 1; Length 913;  
 Best Local Similarity 51.98; Pred. No. 3.79e-01;  
 Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;  
 Db 418 ILTGCLVAIIILLLIALLMLRWLHR 444  
 QY 233 ILISLL-AIIIMVSLLL-LSLWL-WR 256  
 RESULT 8  
 ID LMP1\_EBVR STANDARD; PRT; 386 AA.  
 AC P13198;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-DEC-1992 (Rel. 24, Last annotation update)  
 DE LATENT MEMBRANE PROTEIN 1 (LMP-1) (P63).  
 GN BNLF1.  
 OS Epstein-Barr virus (strain Raji) (Human herpesvirus 4).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae; Lymphocryptovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88219520.  
 RA Hatfull G., Bankier A.T., Barrell B.G., Farrell P.J.;  
 RT "Sequence analysis of Raji Epstein-Barr virus DNA.";  
 RL Virology 164:334-340(1988).  
 CC -!- FUNCTION: THE LATENT MEMBRANE PROTEIN HAS TRANSFORMING ACTIVITY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M20868; AAA66532.1;  
 DR PIR; C28918; LABERJ  
 KW Transmembrane; Phosphorylation; Transforming protein.  
 FT DOMAIN 1 24  
 FT TRANSMEM 25 44  
 FT TRANSMEM 52 72  
 FT TRANSMEM 77 97  
 FT TRANSMEM 105 125  
 FT TRANSMEM 139 159  
 FT TRANSMEM 166 186  
 FT TRANSMEM 187 386  
 FT DOMAIN 187 386  
 FT CYTOPLASMIC.  
 SQ SEQUENCE 386 AA; 42061 MW; 119A03B574455CF4 CRC64;  
 Query Match 36.0%; Score 99; DB 1; Length 386;  
 Best Local Similarity 60.0%; Pred. No. 4.83e-01;  
 Matches 9; Conservative 6; Mismatches 0; Indels 0; Gaps 0;  
 Db 86 ILLMITLLIALLWL 100  
 QY 240 ILLMVSLLLSLWL 254  
 RESULT 9  
 ID ITA6\_HUMAN STANDARD; PRT; 1073 AA.  
 AC P23229; Q14646; Q16508; Q08443;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-FEB-2000 (Rel. 39, Last annotation update)  
 DE INTEGRIN ALPHA-6 PRECURSOR (VLA-6) (CD49F).  
 GN ITGA6.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PANCREAS;  
 RX MEDLINE; 91009492.  
 RA Tamura R.N., Rozzo C., Starr L., Chambers J., Reichardt L.F.,  
 RA Cooper H.M., Quaranta V.;  
 RT "Epithelial integrin alpha 6 beta 4: complete primary structure of  
 RT alpha 6 and variant forms of beta 4.";  
 RL J. Cell Biol. 111:1593-1604(1990).  
 RN [2]  
 RP REVISIONS TO 78 AND 323.  
 RA Quaranta V.;  
 RL Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 7-1073 FROM N.A.  
 RX MEDLINE; 91301154.  
 RA Hogervorst F., Kuikman I., Geurts van Kessel A., Sonnenberg A.;  
 RT "Molecular cloning of the human alpha 6 integrin subunit. Alternative  
 RT splicing of alpha 6 mRNA and chromosomal localization of the alpha 6  
 RT and beta 4 genes.";  
 RL Eur. J. Biochem. 199:425-433(1991).  
 RN [4]  
 RP SEQUENCE OF 709-810 FROM N.A.  
 RX MEDLINE; 93119630.  
 RA Starr L., Quaranta V.;  
 RT "An efficient and reliable method for cloning PCR-amplification  
 RT products: a survey of point mutations in integrin cDNA.";  
 RL Biotechniques 13:612-618(1992).  
 RN [5]  
 RP SEQUENCE OF 927-1073 FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE; 92052235.  
 RA Tamura R.N., Cooper H.M., Collo G., Quaranta V.;  
 RT "Cell type-specific integrin variants with alternative alpha chain  
 RT cytoplasmic domains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10183-10187(1991).  
 RN [6]  
 RP SEQUENCE OF 24-46.  
 RX MEDLINE; 89251596.  
 RA Kajiji S., Tamura R.N., Quaranta V.;



```

FT NP_BIND 209 217 ATP (BY SIMILARITY).
FT BINDING 230 230 ATP (BY SIMILARITY).
FT ACT_SITE 332 332 BY SIMILARITY.
FT CARBOHYD 34 34 POTENTIAL.
FT CARBOHYD 99 99 POTENTIAL.
SQ SEQUENCE 505 AA; 56774 MW; 4C10D09C9105F23C CRC64;

Query Match 34.9%; Score 96; DB 1; Length 505;
Best Local Similarity 44.0%; Pred. No. 9.95e-01;
Matches 11; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

Db 122 LILGPVLALVVALGTLGWRVRR 146
QY 232 FILISSAILLVSVLLSLSLKLR 256

RESULT 11
ID ITB5_PAPCY STANDARD; PRT; 655 AA.
AC Q07441;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE INTEGRIN BETA-5 (FRAGMENT).
GN ITGB5.
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;
OC Papio.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94040831.
RA Shoji M., Hayzer D.J., Kim T.M., Runge M.S., Hanson S.R.;
RT "Human and baboon integrin beta 5 subunit-encoding mRNAs have
RL alternative polyadenylation sites.";
RL Gene 133:307-308(1993).
CC -!- FUNCTION: INTEGRINS ARE A LARGE FAMILY OF CELL SURFACE
CC GLYCOPROTEINS THAT MEDIATE CELL TO CELL & CELL TO MATRIX ADHESION.
CC -!- SUBUNIT: DIMER OF AN ALPHA AND BETA SUBUNIT. BETA-5 ASSOCIATES
CC WITH ALPHA-V.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
CC BONDS.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; L12311; AAA16866.1; -
CC HSP; P04355; 2MRT.
CC PROSITE; PS00243; INTEGRIN_BETA; 2.
CC PROSITE; PS00222; EGF_1; UNKNOWN_2.
CC PROSITE; PS01186; EGF_2; UNKNOWN_2.
CC PFAM; PF00362; Integrin_B; 1.
CC Integrin; Cell adhesion; Transmembrane; Glycoprotein; Repeat;
CC Extracellular matrix; Cytoskeleton.
CC NON_TER 1 575
CC DOMAIN <1 575 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 576 598 POTENTIAL.
CC DOMAIN 599 655 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 321 486 CYSTEINE-RICH REPEATS.
CC REPEAT 321 368 I.
CC REPEAT 369 410 II.
CC REPEAT 411 449 III.
CC REPEAT 450 486 IV.
CC CARBOHYD 203 203 POTENTIAL.
CC CARBOHYD 316 316 POTENTIAL.
CC CARBOHYD 408 408 POTENTIAL.
CC CARBOHYD 442 442 POTENTIAL.

FT NP_BIND 209 217 ATP (BY SIMILARITY).
FT BINDING 230 230 ATP (BY SIMILARITY).
FT ACT_SITE 332 332 BY SIMILARITY.
FT CARBOHYD 34 34 POTENTIAL.
FT CARBOHYD 99 99 POTENTIAL.
SQ SEQUENCE 505 AA; 56774 MW; 4C10D09C9105F23C CRC64;

Query Match 34.9%; Score 96; DB 1; Length 505;
Best Local Similarity 44.0%; Pred. No. 9.95e-01;
Matches 11; Conservative 10; Mismatches 4; Indels 0; Gaps 0;

Db 122 LILGPVLALVVALGTLGWRVRR 146
QY 232 FILISSAILLVSVLLSLSLKLR 256

RESULT 12
ID ITB6_HUMAN STANDARD; PRT; 788 AA.
AC P18564; Q16500;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE INTEGRIN BETA-6 PRECURSOR.
GN ITGB6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-PANCREAS;
RX MEDLINE; 90307659.
RA Sheppard D., Rozzo C., Starr L., Quaranta V., Erle D.J., Pytela R.;
RT "Complete amino acid sequence of a novel integrin beta subunit (beta
RL 6) identified in epithelial cells using the polymerase chain
RL reaction.";
RL J. Biol. Chem. 265:11502-11507(1990).
RN [2]
RP SEQUENCE OF 116-197 FROM N.A.
RX MEDLINE; 93027753.
RA Jiang W.M., Jenkins D., Yuan Q., Leung E., Choo K.H., Watson J.D.,
RA Krissansen G.W.;
RT "The gene organization of the human beta 7 subunit, the common beta
RL subunit of the leukocyte integrins HML-1 and LPAM-1.";
RL Int. Immunol. 4:1031-1040(1992).
CC -!- FUNCTION: INTEGRINS ARE A LARGE FAMILY OF CELL SURFACE
CC GLYCOPROTEINS THAT MEDIATE CELL TO CELL & CELL TO MATRIX ADHESION.
CC -!- SUBUNIT: DIMER OF AN ALPHA AND BETA SUBUNIT. BETA-6 ASSOCIATES
CC WITH ALPHA-V.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
CC BONDS.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; M35198; AAA36122.1; -
CC EMBL; A26609; CAA01832.1; -
CC EMBL; S49380; AAB23690.1; -
CC PIR; A37057; A37057.
CC HSP; P04355; 2MRT.
CC MIM; 147558; -
CC PRINTS; PR01186; INTEGRIN.
CC PROSITE; PS00243; INTEGRIN_BETA; 3.
CC PROSITE; PS00222; EGF_1; UNKNOWN_2.
CC PROSITE; PS01186; EGF_2; UNKNOWN_1.
CC PFAM; PF00362; Integrin_B; 1.
CC Integrin; Cell adhesion; Transmembrane; Glycoprotein; Repeat;
CC Extracellular matrix; Cytoskeleton; Signal.
CC SIGNAL 1 18 POTENTIAL.
CC CHAIN 19 788 INTEGRIN BETA-6.

```







with the Baka/Bakb alloantigen system.";  
 Blood 75:2343-2348(1990).  
 [12] VARIANT GTA ASP-273.  
 RX MEDLINE: 94110321.  
 RA Poncz M., Rifat S., Collier B.S., Newman P.J., Shattil S.J.,  
 RA Parrella T., Fortina P., Bennett J.S.;  
 RA "Glanzmann thrombasthenia secondary to a Gly273->Asp mutation  
 RT adjacent to the first calcium-binding domain of platelet glycoprotein  
 RT IIB.";  
 RL J. Clin. Invest. 93:172-179(1994).  
 [13] VARIANT GTA ASP-449.  
 RX MEDLINE: 94104078.  
 RA Wilcox D.A., Wautier J.-L., Pidard D., Newman P.J.;  
 RA "A single amino acid substitution flanking the fourth calcium binding  
 RT domain of alpha IIB prevents maturation of the alpha IIB beta 3  
 RT integrin complex.";  
 RL J. Biol. Chem. 269:4450-4457(1994).  
 [14] VARIANT GTA HIS-358.  
 RX MEDLINE: 95221604.  
 RA Wilcox D.A., Paddock C.M., Lyman S., Gill J.C., Newman P.J.;  
 RA "Glanzmann thrombasthenia resulting from a single amino acid  
 RT substitution between the second and third calcium-binding domains of  
 RT GPIIb. Role of the GPIIb amino terminus in integrin subunit  
 RT association.";  
 RL J. Clin. Invest. 95:1553-1560(1995).  
 [15] REVIEW ON GTA VARIANTS.  
 RX MEDLINE: 95184171.  
 RA Bray P.P.;  
 RT "Inherited diseases of platelet glycoproteins: considerations for  
 RT rapid molecular characterization.";  
 RL Thromb. Haemost. 72:492-502(1994).  
 [16] VARIANT GTA PRO-778.  
 RX MEDLINE: 98438330.  
 RA Tadokoro S., Tomiyama Y., Honda S., Arai M., Yamamoto N., Shiraga M.,  
 RA Kosugi S., Kanakura Y., Kurata Y., Matsuzawa Y.;  
 RA "A Gln747->Pro substitution in the IIB subunit is responsible for a  
 RT moderate IIBbeta3 deficiency in Glanzmann thrombasthenia.";  
 RL Blood 92:2750-2758(1998).  
 CC -!- FUNCTION: PLATELET MEMBRANE GLYCOPROTEIN IIB-IIIA IS THE PLATELET  
 CC RECEPTOR FOR ADHESIVE PROTEINS. IT BINDS FIBRINOGEN, FIBRONECTIN,  
 CC VITRONECTIN AND VWF. MOST INTEGRINS RECOGNIZE THE SEQUENCE R-G-D  
 CC IN THEIR LIGAND.  
 CC -!- SUBUNIT: DIMER OF AN ALPHA AND BETA CHAINS. THE ALPHA CHAIN IS  
 CC COMPOSED OF AN HEAVY AND LIGHT CHAINS LINKED BY A DISULFIDE BOND.  
 CC ALPHA-IIB ASSOCIATES WITH BETA-3 (GPIIb).  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- POLYMORPHISM: POSITION 874 IS ASSOCIATED WITH PLATELET-SPECIFIC  
 CC ALLOANTIGEN HPA-3/BAK/LEK. HPA-3A/BAK(A)/LEK(A) HAS ILE-874 AND  
 CC HPA-3B/BAK(B)/LEK(B) HAS SER-874. HPA-3B IS INVOLVED IN NEONATAL  
 CC ALLOIMUNE THROMBOCYTOPENIA (NATP).  
 CC -!- DISEASE: DEFECTS IN ITGA2B ARE ONE OF THE CAUSES OF GLANZMANN  
 CC THROMBASTHENIA (GTA), AN AUTOSOMAL RECESSIVE DISORDER WHICH IS THE  
 CC MOST COMMON INHERITED DISEASE OF PLATELETS. GTA IS CHARACTERIZED  
 CC BY LIFELONG MUCOCUTANEOUS BLEEDING OF MILD-TO-MODERATE SEVERITY  
 CC AND THE INABILITY OF THIS INTEGRIN TO RECOGNIZE MACROMOLECULAR OR  
 CC SYNTHETIC PEPTIDE LIGANDS.  
 CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.  
 CC -!- DATABASE: NAME-PROW; NOTE-CD guide CD41 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd41.htm".  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 -----

DR EMBL; J02764; AAA60114.1; -;  
 DR EMBL; M34480; AAA35926.1; -;  
 DR EMBL; M34344; AAA53150.1; -;  
 DR EMBL; M33319; AAA53150.1; JOINED.  
 DR EMBL; M33320; AAA53150.1; JOINED.  
 DR EMBL; X06831; CAA29987.1; -;  
 DR EMBL; M18085; AAA52597.1; -;  
 DR EMBL; M22568; AAA52587.1; -;  
 DR EMBL; M22569; AAA52588.1; -;  
 DR PIR; A28937; A28937.  
 DR PIR; A28581; A28581.  
 DR PIR; A29522; A29522.  
 DR PIR; A32528; A32528.  
 DR PIR; B32528; B32528.  
 DR PIR; S00268; S00268.  
 DR PIR; S05249; S05249.  
 DR HSSP; F11215; 1A8X.  
 DR MIM; 273800; -;  
 DR PRINTS; PRO1185; INTEGRIN.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PFAM; PF00357; integrin\_A; 1.  
 KW Integrin; Cell adhesion; Glycoprotein; Transmembrane; Platelet;  
 KW Extracellular matrix; Cytoskeleton; Signal; Polymorphism;  
 FT SIGNAL 1 31  
 FT CHAIN 32 1039 PLATELET MEMBRANE GLYCOPROTEIN IIB.  
 FT CHAIN 32 902 HEAVY CHAIN.  
 FT CHAIN 903 1039 LIGHT CHAIN.  
 FT DOMAIN 32 993 EXTRACELLULAR.  
 FT TRANSMEM 994 1019 POTENTIAL.  
 FT DOMAIN 1020 1039 CYTOPLASMIC.

Note: remainder of annotations omitted.

Query Match 34.5%; Score 95; DB 1; Length 1039;  
 Best Local Similarity 36.4%; Pred. No. 1.26e+00;  
 Matches 8; Conservative 12; Mismatches 2; Indels 0; Gaps 0;  
 DB 1000 VLGVGLGGLLLTILVLMKV 1021  
 QY 233 ILISLAILLMSLLLSLWKL 254

Search completed: Wed May 10 11:56:14 2000  
 Job time : 101 secs.



\*\*\*\*\*  
 M P S R C H P P  
 \*\*\*\*\*  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 11:56:32 2000; MasPar time 227.94 Seconds

Tabular output not generated. 10.646 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (226-260) from US09376430A.pep (7 of 25)  
 Perfect Score: 275  
 Sequence: 1 KPKLSKFLILSSLAILLMVLSLLSLWKLVRVKKF 35

Scoring table: PAM 150  
 Gap 11

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: sptrembl12  
 1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
 5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organellae  
 9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
 13:sp-vertebrate 14:sp-virus

Statistics: Mean 35.838; Variance 112.182; scale 0.319

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	106	38.5	108	11	ALPHA 7C INTEGRIN (FRA	5.78e-01
2	106	38.5	1106	11	INTEGRIN SUBUNIT ALPHA	5.78e-01
3	106	38.5	1167	4	INTEGRIN ALPHA 7 PRECU	5.78e-01
4	106	38.5	1180	11	INTEGRIN ALPHA 7 PRECU	5.78e-01
5	105	38.2	1135	11	INTEGRIN ALPHA 7 (ALPH	9.04e-01
6	104	37.8	1049	5	F56A8.1 PROTEIN.	1.13e+00
7	103	37.5	787	11	INTEGRIN BETA-6 SUBUNI	1.76e+00
8	101	36.7	371	14	LATENT MEMBRANE PROTEI	1.76e+00
9	101	36.7	381	14	LATENT MEMBRANE PROTEI	1.76e+00
10	101	36.7	418	2	YVFL PROTEIN.	1.76e+00
11	101	36.7	421	2	HYPOTHETICAL 47.3 KD P	1.76e+00
12	100	36.4	169	2	HYPOTHETICAL 18.4 KD P	2.19e+00
13	100	36.4	183	11	RECEPTOR TYROSINE KINA	2.19e+00
14	100	36.4	220	11	RECEPTOR TYROSINE KINA	2.19e+00
15	100	36.4	233	11	INTEGRIN ALPHA IIB SUB	2.19e+00
16	100	36.4	911	11	RECEPTOR-LIKE TYROSINE	2.19e+00
17	99	36.0	371	14	LATENT MEMBRANE PROTEI	2.73e+00
18	99	36.0	382	14	LATENT MEMBRANE PROTEI	2.73e+00
19	99	36.0	382	14	LATENT MEMBRANE PROTEI	2.73e+00
20	99	36.0	386	14	LATENT MEMBRANE PROTEI	2.73e+00

21	98	35.6	81	4	043197	INTEGRIN ALPHA 7A SUBU	3.39e+00
22	98	35.6	148	11	063026	ALPHA 7A INTEGRIN (FRA	3.39e+00
23	98	35.6	1161	11	088731	INTEGRIN ALPHA 7 PRECU	3.39e+00
24	97	35.3	129	11	061741	SERINE/THREONINE KINAS	4.21e+00
25	96	34.9	504	11	063559	TIA5.3 PROTEIN.	5.22e+00
26	95	34.5	355	5	022383	TIA5.3 PROTEIN.	6.48e+00
27	95	34.5	625	5	022384	TIA5.4 PROTEIN.	6.48e+00
28	94	34.2	140	2	005472	ATP SYNTHASE SUBUNIT I	8.02e+00
29	93	33.8	617	10	P93050	ATPK2324.	9.92e+00
30	93	33.8	858	10	Q32776	RECEPTOR PROTEIN KINAS	9.92e+00
31	93	33.8	860	10	Q92NW6	SRK46BRA.	9.92e+00
32	92	33.5	409	15	016268	F40A3.7 PROTEIN.	1.23e+01
33	92	33.5	559	13	093414	ORPHAN G PROTEIN-COUP	1.23e+01
34	92	33.5	614	2	031615	YJBO PROTEIN.	1.23e+01
35	91	33.1	1327	41	015070	KIAA0364.	1.51e+01
36	90	32.7	108	2	085195	COMPETENCE PROTEIN.	1.87e+01
37	90	32.7	353	5	045762	TOSE6.8 PROTEIN.	1.87e+01
38	90	32.7	541	2	046610	CYTOSOLIC C OXIDASE-L	1.87e+01
39	90	32.7	547	2	0927H5	INTEGRAL MEMBRANE PROT	1.87e+01
40	90	32.7	799	11	070308	INTEGRIN BETA 5 (BETA5	1.87e+01
41	90	32.7	799	11	088347	INTEGRIN BETA 5 (BETA5	1.87e+01
42	90	32.7	816	11	070309	INTEGRIN BETA 5 (BETA5	1.87e+01
43	90	32.7	1087	13	Q91778	INTEGRIN ALPHA 6 SUBUN	1.87e+01
44	90	32.7	2410	14	Q9YJW3	POLYPROTEIN.	1.87e+01
45	89	32.4	275	5	Q20587	F49C12.10 PROTEIN.	2.30e+01

ALIGNMENTS

RESULT 1	PRELIMINARY;	PRT;	108 AA.
ID Q63027	01-NOV-1996 (Tremblrel. 01, Created)		
AC Q63027	01-NOV-1996 (Tremblrel. 01, Last sequence update)		
DT 01-NOV-1996	01-NOV-1998 (Tremblrel. 08, Last annotation update)		
DE ALPHA 7C INTEGRIN (FRAGMENT).			
OS Rattus norvegicus (Rat).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
[1]			
RN SEQUENCE FROM N.A.			
RC TISSUE=SKELETAL MUSCLE;			
RX MEDLINE; 94171924.			
RA SONG W.K., WANG W., SAFO H., BIELSER D., KAUFMAN S.;			
RT "Expression of alpha 7 integrin cytoplasmic domains during skeletal			
RT muscle development: alternate forms, conformational change, and			
RT homologues with serine/threonine kinases and tyrosine phosphatases."			
RL J. Cell Sci. 106:1139-1152(1993).			
DR EMBL; X74294; CAA52347.1;			
DR PFAM; PF00357; Integrin_A; 1.			
KW Integrin.			
FT NON_TER.			
SQ SEQUENCE 108 AA; 12016 MW; 1A077241 CRC32;			
Query Match 38.5%; Score 106; DB 11; Length 108;			
Best Local Similarity 59.1%; Pred. No. 5.78e-01;			
Matches 13; Conservative 6; Mismatches 3; Indels 0; Gaps 0;			
Db 71 ILIYAVIAGLVVALLVLLWL 92			
QY 233 ILISSLAILMVLSLLWL 254			
RESULT 2	PRELIMINARY;	PRT;	1106 AA.
ID Q63258	01-NOV-1996 (Tremblrel. 01, Created)		
AC Q63258	01-NOV-1996 (Tremblrel. 01, Last sequence update)		
DT 01-NOV-1996	01-NOV-1999 (Tremblrel. 12, Last annotation update)		
DE INTEGRIN ALPHA CHAIN (FRAGMENT).			
OS Rattus norvegicus (Rat).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			

```

[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=129SV;
RA "SAHER G., ECHTERMEYER F., BEIER D.R., POESCHL E., MAYER U.;
RT "Genomic organization and chromosomal localization of the mouse
RT integrin alpha7 gene.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y1380; CAA73023.1; -.
DR EMBL; Y1383; CAA73023.1; JOINED.
DR EMBL; Y1384; CAA73023.1; JOINED.
DR EMBL; Y1385; CAA73023.1; JOINED.
DR EMBL; Y1386; CAA73023.1; JOINED.
DR EMBL; Y1387; CAA73023.1; JOINED.
DR EMBL; Y1388; CAA73023.1; JOINED.
DR EMBL; Y1389; CAA73023.1; JOINED.
DR EMBL; Y1390; CAA73023.1; JOINED.
DR EMBL; Y1382; CAA73023.1; JOINED.
DR HSSP; P11215; IABX.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PFAM; PF00357; integrin_A_3.
KW Signal; Integrin.
FT SIGNAL 0
SQ SEQUENCE 1180 AA; 129261 MW; 57E6BE73 CRC32;
Query Match 38.5%; # Score 106; DB 11; Length 1180;
Best Local Similarity 59.1%; Pred.No. 5.78e-01;
Matches 13; Conservative 6; Mismatches 3; Indels 0; Gaps
Db 1084 ILJAVLAGLVLLVALLVLWLK 1105
QY 233 ILISSLAILLNVSLILSLWLK 254
RESULT 5 PRELIMINARY; PRT; 1135 AA.
ID Q61738 AC Q61738; P70350;
DT DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE DE 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE DE INTEGRIN ALPHA 7 (ALPHA-7 INTEGRIN).
GN ITGA7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Rutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC [1]
SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=HEART;
RX MEDLINE; 94075378
RA ZIOBER B.L., VU M.P., WALEH N., CRAWFORD J., LIN C.S., KRAMER R.H.;
RT "Alternative extracellular and cytoplasmic domains of the integrin
alpha 7 subunit are differentially expressed during development.";
RJ J. Biol. Chem. 268:26773-26783(1993).
RN [2]
RP SEQUENCE OF 1-70 FROM N.A.
RX MEDLINE; 96394366
RA ZIOBER B.L., KRAMER R.H.;
RT "Identification and characterization of the cell type-specific and
developmentally regulated alpha7 integrin gene promoter.";
RJ J. Biol. Chem. 271:22915-22922(1996).
DR EMBL; L23423; AAA16600.1; -.
DR EMBL; U60419; AAC52772.1; -.
DR HSSP; P11215; IABX.
DR MGD; MGI:102700; Itga7
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PFAM; PF00357; integrin_A_2.
KW Integrin.
SQ SEQUENCE 1135 AA; 124794 MW; F0D9F836 CRC32;
Query Match 38.2%; Score 105; DB 11; Length 1135;
Best Local Similarity 59.1%; Pred.No. 7.23e-01;
Matches 13; Conservative 6; Mismatches 3; Indels 0; Gaps
Db 1039 ILGVLGLVLLVALLVLWLK 1060

```

```

||||: |||: |||: |||: |||: |||:
QY 233 ILISSLAILLMVSLLLSLWKL 254

RESULT 6
ID O45572 PRELIMINARY; PRT; 1049 AA.
AC O45572;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-MAY-1999 (Tremblrel. 06, Last sequence update)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DE F36A8.1 PROTEIN.
GN F36A8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA MCMURRAY A.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISER N., LATREILLE P.,
RA LIGHTING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATKINSON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; 283230; CAB05741.1; -.
SQ SEQUENCE 1049 AA; 121568 MW; 046565CA CRC32;

Query Match 37.8%; Score 104; DB 5; Length 1049;
Best Local Similarity 38.2%; Pred. No. 9.04e-01;
Matches 13; Conservative 15; Mismatches 4; Indels 2; Gaps 2;

Db 417 MASFEVYS-FSMLVVVSLMLVTLTKIWMVNF 449
QY 228 KLSKFLISSLAILLMV-SLLLSLWKLWRVKKF 260

RESULT 7
ID Q920T9 PRELIMINARY; PRT; 787 AA.
AC Q920T9;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE INTEGRIN BETA-6 SUBUNIT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA AREND L.J., SMART A.M., SCHNERMANN J., BRIGGS J.P.;
RA "Mouse beta-6 integrin subunit sequence and expression in adult and
RT developing kidney.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INTEGRINS ARE A LARGE FAMILY OF CELL SURFACE
CC -1- GLYCOPROTEINS THAT MEDIATE CELL TO CELL & CELL TO MATRIX ADHESION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: WITH OTHER BETA CHAINS FROM THE INTEGRIN FAMILY OF
CC CELL-SURFACE RECEPTOR.
DR EMBL; AF115376; AAD17212.1; -.
DR HSSP; P04355; 2MRT.
DR PROSITE; PS00243; INTEGRIN_BETA; 2.
KW Integrin; Cell adhesion; Transmembrane; Glycoprotein; Repeat.
SQ SEQUENCE 787 AA; 86041 MW; 3B65E68A CRC32;

```

```

Query Match 37.5%; Score 103; DB 11; Length 787;
Best Local Similarity 46.4%; Pred. No. 1.13e+00;
Matches 13; Conservative 11; Mismatches 3; Indels 1; Gaps 1;

Db 705 PNIPMIMLVGSLLAILL-IGVVLLCIWKL 731
QY 227 PKLSKFLISSLAILLMVSLLLSLWKL 254

RESULT 8
ID O451E8 PRELIMINARY; PRT; 371 AA.
AC O451E8;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE LATENT MEMBRANE PROTEIN-1.
GN LMP-1.
OS Human herpesvirus 4 (Epstein-Barr virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-AKATA;
RX MEDLINE; 99058177.
RA TORII T., KONISHI K., SAMPLE J., TAKADA K.;
RT "The truncated form of the Epstein-Barr virus LMP-1 is dispensable or
RT complementable by the full-length form in virus infection and
RT replication.";
RL Virology 251:273-278(1998).
DR EMBL; AF093534; AAD19837.1; -.
SQ SEQUENCE 371 AA; 40391 MW; B05F9B66 CRC32;

Query Match 36.7%; Score 101; DB 14; Length 371;
Best Local Similarity 47.6%; Pred. No. 1.76e+00;
Matches 10; Conservative 10; Mismatches 0; Indels 1; Gaps 1;

Db 80 LGGGLGLLLMTLLIALWNL 100
QY 235 ISSLAILL-MVSLLLSLWKL 254

RESULT 9
ID Q89558 PRELIMINARY; PRT; 381 AA.
AC Q89558;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE LATENT MEMBRANE PROTEIN.
GN LMP 1.
OS Human herpesvirus 4 (Epstein-Barr virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 9304689.
RA CHEN M.L., TSAI C.N., LIANG C.L., SHU C.H., HUANG C.R., SULITZANU D.,
RA LIU S.T., CHANG Y.S.;
RT "Cloning and characterization of the latent membrane protein (LMP) of
RT a specific Epstein-Barr virus variant derived from the nasopharyngeal
RT carcinoma in the Taiwanese population.";
RL Oncogene 7:2131-2140(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA CHANG Y.S., CHEN M.L., TSAI C.N., HUANG C.R., LIANG C.L., SHU C.H.,
RA SULITZANU D., LIU S.T.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; X66863; CAA47332.1; -.
DR EMBL; M83720; AAA45886.1; -.
KW Membrane.
SQ SEQUENCE 381 AA; 41372 MW; 470FAA99 CRC32;

Query Match 36.7%; Score 101; DB 14; Length 381;

```

Best Local Similarity 47.6%; Pred. No. 1.76e+00;  
Matches 10; Conservative 10; Mismatches 0; Indels 1; Gaps 1;

```
Db      80  LGGLGILLMITILLIALWNL 100
        ::::| |::| |::| |:|
QY      235  ISSLAILL-MVSLLLLSLWKL 254
```

RESULT	10		
ID	O32261	PRELIMINARY;	PRT; 418 AA.
AC	O32261;		
DT	01-JAN-1998 (TReMBLrel. 05, Created)		
DT	01-JAN-1998 (TReMBLrel. 05, Last sequence update)		
DT	01-NOV-1999 (TReMBLrel. 12, Last annotation update)		
DE	YVFL PROTEIN.		
GN	YVFL.		
OS	Bacillus subtilis.		
OC	Bacteria; Firmicutes.		
OC	Bacillus/Clostridium group;		
OC	Bacillus/Staphylococcus group; Bacillus.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-168;		
RX	MEDLINE: 98044033.		

MEDLINE: 98044033.  
 KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
 RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,  
 RA ARISSO R., BOURIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,  
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPURANO V., CARTER N.M.,  
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,  
 RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMERSON P.T.,  
 RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,  
 RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,  
 RA GHIM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,  
 RA GUISEPPI G., GUY B.J., HAGA K., HALECH J., HARWOOD C.R., HENAUT A.,  
 RA HILBERT H., HOLSAPPEL S., HOSONO S., HULZO M.F., IATA M., JONES L.,  
 RA JORIS B., KARAMATA D., KASHARA Y., KLAERR-BLANCHARD M., KLEIN C.,  
 RA KOBAYASHI Y., KOTTER P., KONINGSTEIN G., KROGH S., KUMANO M.,  
 RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,  
 RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,  
 RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,  
 RA NOONE D., O'REILLY M., OGAWA K., OGIWARA K., OUDEGA B., PARK S.H.,  
 RA PARRO V., POHL T.M., PORTETELLE D., PORWOLLIK S., PRESCOTT A.M.,  
 RA PRESECAN E., PUTIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,  
 RA RIEGER M., REVOLTA C., ROCHA E., ROCHE B., ROSE M., SADAAE Y.,  
 RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOFFONE F.,  
 RA SEKIGUCHI J., SEKOWSKA A., SERO S.J., SERROR P., SHIN B.S.,  
 RA SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,  
 RA TAREUCHI M., TANAKOSHI A., TANAKA T., TERPSTRA P., TOGNONI A.,  
 RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,  
 RA VIARI A., WAMBUUT R., WEDLER E., WEDLER H., WEITENEGGER T.,  
 RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,  
 RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.,  
 RA "The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*,"  
 Nature 390:249-256(1997).

[2]  
SEQUENCE FROM N.A.  
STRAIN=168;  
KUNST F.; OGASAWARA N., YOSHIKAWA H., DANCHIN A.;  
Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
EMBL; Z99121; CAB15420.1; -;  
PROSITE: PS00402; BPD\_TRANSP\_INN\_MEMBER; 1.  
PFAM; PF00528; BPD\_transp; 1.  
transport; transmembrane.  
SEQUENCE 418 AA; 46978 MW; 2F0E779A CRC32;

```
Query Match      36.7%; Score 101; DB 2; Length 418;
Best Local Similarity 35.5%; Pred. No. 1.76e+00;
Matches 11: Conservative 11; Mismatches 0; Indels 0;
```

	12, conserved	17, conserved	5, indels	0, gaps
bb	384	SOYSLAALATILLSVFVISIALMQFROTKSF	414	
	:   :   :   :   :   :   :   :			
yy	230	SKFTILSSLAILLMVSLLLSLWKLWVKKF	260	

```

RESULT      11
ID          CO07010      PRELIMINARY;      PRT;      421 AA.
AC          CO07010;
DT          01-JUL-1997 (TReMBLrel. 04, Created)
DT          01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT          01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE          HYPOTHETICAL 47.3 KD PROTEIN.
DE          YVEL.
DE          GN
OS          Bacillus subtilis.
OC          Bacteria; Firmicutes; Bacillus/Clostridium group;
OC          Bacillus/Staphylococcus group; Bacillus.
RN          [1]
RP          SEQUENCE FROM N.A.
RR          DENIZOT F.C.;
RA          Submitted (APR-1997) to the EMBL/GenBank/DBDJ databases.
RA          EMBL; 294043; CAB08006.1; -.
DR          PROSITE; PS00402; BPD_TRANSF_INN_MEMBR; 1.
DR          PFAM; PF00528; BPD_transp; 1.
DR          Hypothetical protein; Transport; Transmembrane.
SK          SEQUENCE 421 AA; 47306 MW; 136CA72E CRC32;
SQ

```

```

Query Match      36.7%;   Score 101;   DB 2;   Length 421;
Best Local Similarity 35.5%;   Pred. No. 1.76e+00;
Matches 11;   Conservative 11;   Mismatches 9;   Indels 0;   Gaps:
db 387  SOYSAALITLLSVFVSIALMQPQTKSF 417
      | - - - - - | | | | | | | | | |
QY 230  SRFILISSLAILLMVSLLLSLKLVRYKKE 260
      | | | | | | | | | | | | | | |

```

RESULT	12
IID	Q55628
AC	Q55628;
DDT	01-NOV-1996 (TrEMBLrel. 01, Created)
DDT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DDT	01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DDE	HYPOTHETICAL 18.4 KD PROTEIN.
	YCF22.
OOS	Synechocystis sp. (strain PCC 6803).
OOC	Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN	[1]
RRP	SEQUENCE FROM N.A.
RCR	STRAIN-PCC6803;
ERRA	TABATA S.;
RL	Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.

[2]  
SEQUENCE FROM N.A.  
STRAIN-PCC6803;  
MEDLINE: 96127529.  
KANEKO T., TANAKA A., SATO S., KOTANI H., SAZUKA T., MIYAJIMA N.,  
SUGIURA M., TABATA S.;  
"Sequence analysis of the genome of the unicellular cyanobacterium  
Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
region from map positions 64% to 92% of the genome.";  
DNA Res. 2:153-166(1995).  
[3]  
SEQUENCE FROM N.A.  
STRAIN-PCC6803;  
MEDLINE: 97061201.  
KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,  
MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,  
HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NAKRO K., OKUMURA  
SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,  
TABATA S.;  
"Sequence analysis of the genome of the unicellular cyanobacterium  
Synechocystis sp. strain PCC6803. II. Sequence determination of the  
entire genome and assignment of potential protein-coding regions.";  
DNA Res. 3:109-136(1996).  
EMBL: D64000; BAA10141.1.  
Hypothetical protein.  
SEQUENCE 169 AA; 18358 MW; E3C768B4 CRC32;

US-09-376-430-2-07.rspt

Thu May 11 06:49:40 2000

```

Query Match          36.4%; Score 100; DB 2; Length 169;
Best Local Similarity 33.3%; Pred. No. 2.19e+00;
Matches 12; Conservative 14; Mismatches 9; Indels 1; Gaps 1;

Db 6 NORLSPVLQSSGILLLTALILIAFSTLMLKNIRF 41
QY 226 KPLKSLKILSSLAILLMVSLLSLLSLWKLW-RVKKF 260

RESULT 13
ID Q64107 PRELIMINARY; PRT; 183 AA.
AC Q64107; 1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE RECEPTOR TYROSINE KINASE (FRAGMENT).
GN PTK-38.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 95320273.
RA SAKUMA S., SAYA H., IJICHI A., TOFILON P.J.;
RT "Radiation induction of the receptor tyrosine kinase gene Ptk-3 in
normal rat astrocytes.";
RL Radiat. Res. 143:1-7(1995).
DR EMBL; S77556; AAB34729.1; -.
FT NON_TER 1
SQ SEQUENCE 183 AA; 19717 MW; 0BIAD57B CRC32;

Query Match          36.4%; Score 100; DB 11; Length 183;
Best Local Similarity 51.9%; Pred. No. 2.19e+00;
Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;

Db 15 ILIGCLVAIIILLIIILMLWRLHWR 41
QY 233 ILISSL-AILLMVSLLSLWKLW-R 256

RESULT 14
ID Q64108 PRELIMINARY; PRT; 220 AA.
AC Q64108;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE RECEPTOR TYROSINE KINASE (FRAGMENT).
GN PTK-3L.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 95320273.
RA SAKUMA S., SAYA H., IJICHI A., TOFILON P.J.;
RT "Radiation induction of the receptor tyrosine kinase gene Ptk-3 in
normal rat astrocytes.";
RL Radiat. Res. 143:1-7(1995).
DR EMBL; S77556; AAB34729.1; -.
FT NON_TER 1
SQ SEQUENCE 220 AA; 23670 MW; 8D79D0DF CRC32;

Query Match          36.4%; Score 100; DB 11; Length 220;
Best Local Similarity 51.9%; Pred. No. 2.19e+00;
Matches 14; Conservative 7; Mismatches 3; Indels 3; Gaps 3;

Db 15 ILIGCLVAIIILLIIILMLWRLHWR 41
QY 233 ILISSL-AILLMVSLLSLWKLW-R 256

RESULT 15
ID Q9Z2M0 PRELIMINARY; PRT; 233 AA.

```

```

Q9Z2M0;
AC 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE INTEGRIN ALPHA IIB SUBUNIT (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 95320273.
RA ROUT U.K., ARMANT D.R.;
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF045019; AAD02339.1; -.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
KW Integrin.
FT NON_TER 1
SQ SEQUENCE 233 AA; 26061 MW; F8A1D2DB CRC32;

Query Match          36.4%; Score 100; DB 11; Length 233;
Best Local Similarity 42.9%; Pred. No. 2.19e+00;
Matches 9; Conservative 10; Mismatches 2; Indels 0; Gaps 0;

Db 206 VLVGVGLGILLTLIVLAWK 226
QY 233 ILISSLAILLMVSLLSLLWKLW 253

Search completed: Wed May 10 12:00:30 2000
Job time : 238 secs.

```





\*\*\*\*\*

MP5RELH (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 12:09:29 2000; MasPar time 2.81 Seconds  
Tabular output not generated. 67.401 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (261-268) from US09376430A.pap (8 of 25)  
Perfect Score: 56  
Sequence: 1 LIPSVDP 8

Scoring table:

PAM 150

Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: a-geneseq35

1:geneseqp

Statistics: Mean 15.198; Variance 46.309; scale 0.328

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	ID	Description	Pred. No.
1	45	80.4	W42448	C. heterostrophus mati	1.89e+02
2	44	78.6	R80575	Mature Bone Morphogeni	2.41e+02
3	44	78.6	W06539	Human bone morphogenetic	2.41e+02
4	44	78.6	R47586	Human bone formation-i	2.41e+02
5	43	76.8	R22983	Canine herpesvirus gly	3.08e+02
6	43	76.8	W72850	Canine herpes virus pr	3.08e+02
7	43	76.8	W22958	Canine herpesvirus gly	3.08e+02
8	43	76.8	R77026	Canine herpesvirus 9D	3.08e+02
9	43	76.8	W72858	Canine herpes virus pr	3.08e+02
10	43	76.8	W04723	Aromatic acyl transfer	3.08e+02
11	42	75.0	Y11574	Human 5' EST secreted	3.92e+02
12	42	75.0	R47587	Rat bone formation-ind	3.92e+02
13	42	75.0	R74970	Murine growth differen	3.92e+02
14	42	75.0	W71468	Cercospora nicotianae	3.92e+02
15	42	75.0	W48898	Candida albicans CaBem	3.92e+02
16	41	73.2	R51651	Osteogenic unglycosyla	4.99e+02
17	41	73.2	W03217	BMP-3 fragment 80-98	4.99e+02
18	41	73.2	R06883	HTLV-I gag 392-411 dia	4.99e+02
19	41	73.2	Y11554	Human 5' EST secreted	4.99e+02
20	41	73.2	W30345	Fragment of bone morph	4.99e+02
21	41	73.2	W71050	Multiple sclerosis ass	4.99e+02
22	41	73.2	W00233	Human BMP3(fx) morphog	4.99e+02
23	41	73.2	W36865	Conserved 7 cysteine m	4.99e+02

## ALIGNMENTS

RESULT 1

ID W42448 standard; protein: 343 AA.

AC W42448;

DT 10-JUN-1998 (first entry)

DE C. heterostrophus mating factor MAT1-2 used in a transgenic plant.

KW Fungal pheromone; mating factor; transgenic plant; fungal infection;

KW Cochliobolus heterostrophus.

OS Cochliobolus heterostrophus.

PN W09748719-A1.

PD 24-DEC-1997.

PF 17-JUN-1997: U10364.

PR 17-JUN-1996: US-019598.

PA (TEXA) UNIV TEXAS A & M SYSTEM.

PI Beckerman JL, Ebbola DJ, Kazmierczak P, McCabe P,

PI Van Alfen NJ, Zhang LJ;

DR WPI: 98-06307/06.

PT Transgenic plant, whose genome contains a fungal pheromone gene

PT has increased resistance to fungal infections

PS Example 5; Page 98; 141pp; English.

CC This is a mating factor MAT1-2 from Cochliobolus heterostrophus used

CC in the preparation of a transgenic plant whose genome is augmented by the

CC introduction of a fungal pheromone gene, which confers resistance to

CC fungal infections. The transgenic plant is prepared by a process which

CC comprises obtaining a polynucleotide composition comprising 1 or more

CC fungal pheromone genes, contacting recipient plant cells with the

CC polynucleotide composition, regenerating plants from recipient cells

CC which have received the polynucleotide composition and identifying a

CC fertile transgenic plant whose genome has been augmented relative to

CC that of the corresponding non-transgenic recipient cells through the

CC stable introduction of the polynucleotide composition. The invention

CC provides for the identification of DNA segments encoding the fungal

CC pheromones, and compositions comprising pheromone polypeptides, which

CC inhibit fungal growth and development and prevent or reduce fungal

CC infection and/or disease in plant host cells. Also provided are methods

CC and processes utilising fungal pheromone and pheromone derivatives in the

CC prevention and treatment of fungal colonisation of a plant.

SQ Sequence: 343 AA;

Query Match: 80.4%; Score 45; DB 1; Length 343;

Best Local Similarity 71.4%; Pred. No. 1.89e+02;

Matches: 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 237 LIPTYPE 243

QY 261 LIPSVDP 267

24	41	73.2	104	1	R57982	Human BMP3(fx).	4.99e+02
25	41	73.2	104	1	W40196	Human partial BMP-3 pr	4.99e+02
26	41	73.2	104	1	R60977	Human BMP3(fx).	4.99e+02
27	41	73.2	104	1	R33414	Human BMP3(fx) seven c	4.99e+02
28	41	73.2	183	1	W84211	Bone morphogenetic pro	4.99e+02
29	41	73.2	245	1	R36739	Human bone morphogenet	4.99e+02
30	41	73.2	223	1	W15759	Cotton fibrous tissue	4.99e+02
31	41	73.2	386	1	W71049	Multiple sclerosis ass	4.99e+02
32	41	73.2	386	1	W12370	Partial sequence of MS	4.99e+02
33	41	73.2	398	1	R22212	Sequence of interleuki	4.99e+02
34	41	73.2	415	1	R22211	Sequence of interleuki	4.99e+02
35	41	73.2	415	1	R22217	Sequence of p24 flanke	4.99e+02
36	41	73.2	429	1	P50232	Human T-cell leukemia	4.99e+02
37	41	73.2	430	1	R03724	Ehrlichia sp. W20.2 pr	4.99e+02
38	41	73.2	514	1	W81976	B.diminuta pimelyi COA	4.99e+02
39	41	73.2	525	1	W99453	Protein pMON32390.pap	4.99e+02
40	41	73.2	749	1	W78002	Granulocyte colony sti	4.99e+02
41	41	73.2	783	1	R11741	MSRV-1 pol protein seq	4.99e+02
42	41	73.2	796	1	W36024	Multiple sclerosis ass	4.99e+02
43	41	73.2	796	1	W71080	Human granulocyte colo	4.99e+02
44	41	73.2	800	1	W10486	Clone 25-1 encoded hum	4.99e+02
45	41	73.2	801	1	R11742		

RESULT 2  
 ID R80575 standard; Protein; 119 AA.  
 AC R80575;  
 DT 18-APR-1996 (first entry)  
 DE Mature Bone Morphogenic Protein-10.  
 KW BMP-10; bone morphogenic protein-10; osteoporosis; lung cancer;  
 treatment; de novo bone formation.  
 OS Homo sapiens.  
 FH key Location/Qualifiers  
 FT domain 61..119  
 FT /note= "active domain"  
 PN WO9524474-A1.  
 PD 14-SEP-1995.  
 PF 12-MAY-1994; U05292.  
 PR 10-MAR-1994; US-209214.  
 PI (HUMA-) HUMAN GENOME SCI INC.  
 PI Coleman TA, Hastings GA, He WW, Rosen CA, Wuhe W;  
 DR WPI, 95-328267/42.  
 DR N-PSDB; Q98822.  
 PT Human bone morphogenic protein (BMP)-10 polypeptide - is used for de  
 novo bone formation e.g. during surgical insertion of protheses  
 PS Claim 1; Fig 1A; 35pp; English.  
 CC The mature human bone morphogenic protein (BMP)-10 is encoded by Q98822.  
 CC residues 414 to 472. It was shown that the mRNA for BMP-10 is abundant  
 in lung, and anti-BMP-10 antibodies are useful in an immunoassay to  
 detect lung abnormalities, esp. lung cancer where there is an increase  
 in BMP-10 expression. Human BMP-10 is useful for inducing de novo bone  
 formation, e.g. during surgical insertion of protheses, for the  
 CC treatment of non-union bone fractures or osteoporosis.  
 SQ Sequence 119 AA;  
 Query Match 78.6%; Score 44; DB 1; Length 119;  
 Best Local Similarity 50.0%; Pred. No. 2.41e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Db 75 IIPGIPEP 82  
 QY 261 LIPSVDPDP 268  
 RESULT 3  
 ID W06539 standard; Protein; 478 AA.  
 AC W06539;  
 DT 12-MAR-1997 (first entry)  
 DE Human bone morphogenic protein-10.  
 KW Bone morphogenic protein 10; BMP-10; fracture; osteoporosis;  
 periodontal disease; diagnosis; therapy;  
 transforming growth factor beta.  
 OS Homo sapiens.  
 FH key Location/Qualifiers  
 FT region 344..478  
 FT /note= "amino acids 344-478 constitute the active  
 portion of BMP-10"  
 PN WO9639431-A1.  
 PD 12-DEC-1996.  
 PF 06-JUN-1995; U07915.  
 PR 06-JUN-1995; WO-007915.  
 PI (HUMA-) HUMAN GENOME SCI INC.  
 PI Coleman TA, Hastings GA, He WW, Rosen CA;  
 DR WPI; 97-043066/04.  
 DR N-PSDB; T45868.  
 PT Isolated human bone morphogenic protein-10 and corresponding DNA -  
 can be used to stimulate bone formation and to develop prods. for  
 diagnosis and therapy  
 PS Claim 11; Fig 1; 49pp; English.  
 CC Human morphogenic protein 10 (BMP-10) (W06539) is a novel member of  
 the transforming growth factor beta superfamily. Its amino acid  
 sequence was deduced from a cDNA clone (T45868) obtd. from a foetal  
 lung cDNA library. BMP-10 polypeptides, esp. the active fragment  
 comprising amino acids 344-478, can be produced in transformed  
 CC host cells (e.g. E. coli, CHO). They can be used to promote de

CC novo bone formation, to treat bone fractures, osteoporosis and  
 periodontal disease, and to raise diagnostic antibodies.  
 SQ Sequence 478 AA;  
 Query Match 78.6%; Score 44; DB 1; Length 478;  
 Best Local Similarity 50.0%; Pred. No. 2.41e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Db 434 IIPGIPEP 441  
 QY 261 LIPSVDPDP 268  
 RESULT 4  
 ID R47586 standard; Protein; 478 AA.  
 AC R47586;  
 DT 18-JUL-1994 (first entry)  
 DE Human bone formation-inducing protein.  
 KW Bone formation-inducing protein; BIP; human; rat; osteoporosis;  
 bone deficiency; alveolar pyorrhae; fracture; vertebrate.  
 OS Homo sapiens.  
 FH key Location/Qualifiers  
 FT peptide 1..368  
 FT /label= sig\_peptide  
 FT protein 369..478  
 FT /label= mat\_protein  
 FT /note= "claim 1, 2"  
 FT peptide 1..367  
 FT /label= sig\_peptide  
 FT protein 368..478  
 FT /label= mat\_protein  
 FT peptide 1..366  
 FT /label= sig\_peptide  
 FT protein 367..478  
 FT /label= mat\_protein  
 PN WO9401557-A.  
 PD 20-JAN-1994.  
 PF 09-JUL-1993; J00952.  
 PR 13-JUL-1992; JP-206996.  
 PA (SUMO) SUMITOMO METAL IND LTD.  
 PI Fukuda K, Hino J, Kangawa K, Keshita N, Konno Y;  
 PI Takao M.  
 DR WPI; 94-035064/04.  
 DR N-PSDB; Q54050.  
 PT Bone formation-inducing protein - for therapy of diseases  
 PT involving osteoporosis, a bone deficiency such as alveolar  
 pyorrhae etc. and bone fracture  
 PS Disclosure; Page 40-43; 57pp; English.  
 CC Protein having improved bone formation inducing-activity  
 CC has been provided. BIP mRNA may be obtained from the tissue  
 CC of a vertebrate (e.g. human, rat) and used in recombinant  
 CC DNA techniques for the prodn. of the protein. The BIP  
 CC is useful in pharmaceuticals.  
 SQ Sequence 478 AA;  
 Query Match 78.6%; Score 44; DB 1; Length 478;  
 Best Local Similarity 50.0%; Pred. No. 2.41e+02;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Db 434 IIPGIPEP 441  
 QY 261 LIPSVDPDP 268  
 RESULT 5  
 ID W22983 standard; Protein; 119 AA.  
 AC W22983;  
 DT 20-FEB-1998 (first entry)  
 DE Canine herpesvirus glycoprotein D (CgD) partial sequence.  
 KW Vaccine, vector; gene therapy; canid; dog; CHV; glycoprotein D;  
 CgD.  
 OS Canine herpesvirus.  
 PN WO9729772-A1.

US-09-376-430-2-08.rag

Thu May 11 06:49:41 2000

```

PD 21-AUG-1997.
PF 14-FEB-1997; U04115.
PR 15-FEB-1996; US-602010.
PA (HESK-) HESKA CORP.
PI Frank RA, Haanes EJ;
DR WPI: 97-424758/39.
DR N-PSDB: T75616.
DR Recombinant canine herpes virus and its genome - useful as vaccine
PT to protect canids against infectious, metabolic or genetic diseases
PS Claim 57; Page 143-144; 240pp; English.
CC This protein comprises a portion of glycoprotein D (Cgd) of canine
CC herpesvirus (CHV). Its sequence was deduced from a coding region
CC found in CHV genomic DNA molecules NCUS5495 (T75590) and NCUS10592
CC (T75616). Cgd (see also W22998) can be expressed in transformed
CC host cells. CHV proteins such as Cgd, nucleic acids, and
CC antibodies raised against CHV proteins, can be used to protect
CC canids against CHV infection. Novel recombinant CHV and novel
CC recombinant CHV genomes that contain heterologous nucleic acid
CC molecules inserted e.g. into the Cgd gene, can be used as vaccines
CC to protect canids against infectious, metabolic or genetic diseases.
CC Sequence 119 AA;
SQ

Query Match 76.8%; Score 43; DB 1; Length 119;
Best Local Similarity 50.0%; Pred. No. 3.08e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 33 MIPDIPNP 40
QY 261 LIPSVDPF 268

RESULT 6
ID W72650 standard; Protein; 119 AA.
AC W72650;
DE 07-JAN-1999 (first entry)
DE Canine herpes virus protein sequence PCgD-119.
DE Canine herpes virus; CHV; recombinant canine herpes virus vector;
KW genome; vaccination; dog; protozoan; helminth; ectoparasite; bacteria;
KW virus infection.
OS Canine herpes virus.
PN US804197-A.
DR 08-SEP-1998. 680726.
PF 12-JUL-1996; 680726.
PR 12-JUL-1996; US-680726.
PR 15-FEB-1996; US-602010.
PA (HESK-) HESKA CORP.
PI Frank RS, Haanes EJ;
DR WPI: 98-505590/43.
DR N-PSDB: V66906.
DR Canine herpes virus nucleic acids - useful for producing recombinant
PT canine herpes virus vectors
PS Claim 7; Column 77-78; 103pp; English.
CC The present sequence represents a canine herpes virus (CHV) protein
CC sequence. The nucleic acid sequence which encodes the protein can be
CC used for DNA vaccination of dogs against CHV and also other infective
CC agents such as protozoans, helminths, ectoparasites, bacteria and
CC viruses. CHV can be formulated by incorporation of heterologous nucleic
CC acid molecules as a single multivalent therapeutic composition against
CC a variety of canine pathogens.
CC Sequence 119 AA;
SQ

Query Match 76.8%; Score 43; DB 1; Length 119;
Best Local Similarity 50.0%; Pred. No. 3.08e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 33 MIPDIPNP 40
QY 261 LIPSVDPF 268

RESULT 7
ID W22998 standard; Protein; 345 AA.
AC W22998;

```

```

DT 20-FEB-1998 (first entry)
DE Canine herpesvirus glycoprotein D PCgD345.
KW Vaccine; vector; gene therapy; canid; dog; CHV; Cgd; PCgD345;
OS glycoprotein D.
OS Canine herpesvirus.
PN WO9729772-A1.
DR WPI: 97-424758/39.
DR N-PSDB: T75616.
DR Recombinant canine herpes virus and its genome - useful as vaccine
PT to protect canids against infectious, metabolic or genetic diseases
PS Example 10; Page 177-178; 240pp; English.
CC This protein comprises canine herpesvirus (CHV) glycoprotein D
CC (PCgD345). Its sequence was deduced from a coding region found in
CC CHV genomic DNA molecule NCUS10592 (see T75616). PCgD345 can be
CC expressed in transformed host cells. CHV proteins, nucleic acids,
CC and antibodies raised against CHV proteins, can be used to protect
CC canids against CHV infection. Novel recombinant CHV and novel
CC recombinant CHV genomes that contain heterologous nucleic acid
CC molecules inserted e.g. into Cgd gene, can be used as vaccines to
CC protect canids against infectious, metabolic or genetic diseases.
CC Sequence 345 AA;
SQ

Query Match 76.8%; Score 43; DB 1; Length 345;
Best Local Similarity 50.0%; Pred. No. 3.08e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 33 MIPDIPNP 40
QY 261 LIPSVDPF 268

RESULT 8
ID R77026 standard; Protein; 345 AA.
AC R77026;
DE 20-APR-1996 (first entry)
DE Canine herpesvirus gd homologue.
DE Canine herpesvirus gd; vector; attenuation; poxvirus; vaccinia virus;
KW CHV; glycoprotein gd; canarypox virus; ALVAC1 VCP294; dog; puppy; immunisation; antigen;
KW vaccine.
OS Canine herpesvirus.
PN Key
DR WPI: 98-505590/43.
DR N-PSDB: T01404.
DR Nucleic acids encoding canine herpes virus (CHV) gB, gC and gD
PT glycoproteins - also glyco:proteins and vectors, for the
PT immunisation of neonatal puppies and adult dogs against CHV
PS Example 11; Fig. 7A-C; 241pp; English.
CC A canine herpesvirus (CHV) protein has a predicted amino acid
CC sequence (R77026) that shows significant homology with the gD
CC glycoprotein of herpesviruses. It is the product of a gene (see
CC T01404) isolated by cloning of CHV genomic DNA fragments in
CC pBluescriptSK. CHV glycoprotein gB (R77024) and gC (R77025)
CC homologues have also been obtd. These glycoproteins, including
CC recombinant glycoproteins expressed from attenuated recombinant
CC virus vectors, e.g. ALVAC recombinant VCP294 (see T01468), can be
CC used in antigenic, immunological or vaccine compositions to protect
CC puppies and adult dogs against CHV.

```

```

SQ Sequence 345 AA;
Query Match 76.8%; Score 43; DB 1; Length 345;
Best Local Similarity 50.0%; Pred. No. 3.08e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 33 MIPDIPNP 40
Qy 261 LIPSVDPDP 268
||| :|||

RESULT 9
ID W72658 standard; Protein; 345 AA.
AC W72658;
DE 07-JAN-1999 (first entry)
DE Canine herpes virus protein sequence PCGP-345.
KW Canine herpes virus; CHV; recombinant canine herpes virus vector;
KW genome; vaccination; dog; protozoan; helminth; ectoparasite; bacteria;
KW virus infection.
OS Canine herpes virus.
PN US804197-A.
PD 08-SEP-1998.
PF 12-JUL-1996; 680726.
PR 12-JUL-1996; US-680726.
PR 15-FEB-1996; US-602010.
PA (HESK-) HESKA CORP.
PI Frank RS, Haanes EJ;
DR WPI; 98-505590/43.
DR N-PSDB; V66940.
PT Canine herpes virus nucleic acids - useful for producing recombinant
PT Canine herpes virus vectors
PS Claim 7: Column 129-130; 103pp; English.
CC The present sequence represents a canine herpes virus (CHV) protein
CC sequence. The nucleic acid sequence which encodes the protein can be
CC used for DNA vaccination of dogs against CHV and also other infective
CC agents such as protozoans, helminths, ectoparasites, bacteria and
CC viruses. CHV can be formulated by incorporation of heterologous nucleic
CC acid molecules as a single multivalent therapeutic composition against
CC a variety of canine pathogens.
SQ Sequence 345 AA;

Query Match 76.8%; Score 43; DB 1; Length 345;
Best Local Similarity 50.0%; Pred. No. 3.08e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 33 MIPDIPNP 40
Qy 261 LIPSVDPDP 268
||| :|||

RESULT 10
ID W04723 standard; Protein; 479 AA.
AC W04723;
DE 06-FEB-1997 (first entry)
DE Aromatic acyl transferase of Gentiana triflora var. japonica.
KW Aromatic acyl transferase; transformation; anthocyanin pigment;
KW plants; acylation; colour; tone; colouration; colour change;
KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;
KW Scenecio cruentus; Lavandula angustifolia.
OS Gentiana triflora var. japonica (Clone pGAT106).
PN W09625500-A1.
PD 22-AUG-1996.
PF 16-FEB-1996; J00348.
PR 17-FEB-1995; JP-067159.
PR 29-JUN-1995; JP-196915.
PR 30-JAN-1996; JP-046534.
PA (SUNK) SUNTORY LTD.
PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
PI Nakao M, Tanaka Y, Yonekura K;
DR WPI; 96-393401/39.
DR N-PSDB; T37309.
PT DNA coding for aromatic acyl transferase - for transforming plants
PT which produce anthocyanin pigments and thus altering colour tone,

SQ Sequence 345 AA;
Query Match 76.8%; Score 43; DB 1; Length 345;
Best Local Similarity 50.0%; Pred. No. 3.08e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 33 MIPDIPNP 40
Qy 261 LIPSVDPDP 268
||| :|||

RESULT 11
ID Y11574 standard; Protein; 150 AA.
AC Y11574;
DE 16-JUN-1999 (first entry)
DE Human 5' EST secreted protein SEQ ID NO:226.
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
OS Homo sapiens.
PN W09506439-A2.
PD 11-FEB-1999.
PF 31-JUL-1998; IB1233.
PR 01-AUG-1997; US-904468.
PA (GEST) GENSET.
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
DR WPI; 99-153700/13.
DR N-PSDB; X40292.
PT New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries derived from liver, lung, large intestine, colon,
PT thyroid and pancreas tissue
PS Claim 27; Page 332-333; 398pp; English.
CC X40251 to X40397 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins, and encode the proteins given in Y11533 to Y11679,
CC respectively. The proteins given represent the signal peptide and an
CC N-terminal fragment of a secreted protein. The nucleic acid sequences
CC can be used for producing secreted human gene products. The proteins
CC can be used to develop products for diagnosis and therapy. The proteins
CC obtained may have cytokine activity, cell proliferation/differentiation
CC activity, haematopoiesis regulating activity, tissue growth regulating
CC activity, reproductive hormone regulating activity, chemotactic/
CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC ligand activity, anti-inflammatory activity, tumour inhibition activity
CC or other activities. The products can be used in forensic, gene therapy
CC and chromosome mapping procedures. The sequences can also be used for
CC obtaining corresponding promoter sequences. The nucleic acids encoding
CC the signal peptide can be used for directing extracellular secretion of
CC a polypeptide or the insertion of a polypeptide into a membrane, or
CC importing a polypeptide into a cell.
SQ Sequence 150 AA;

Query Match 75.0%; Score 42; DB 1; Length 150;
Best Local Similarity 62.5%; Pred. No. 3.92e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 31 LIPAPDPP 38
Qy 261 LIPSVDPDP 268
||| :|||

RESULT 12

```

ID R47587 standard; Protein; 476 AA.  
AC 18-JUL-1994 (first entry)  
DE Rat bone formation-inducing protein.  
KW Bone formation-inducing protein; BIP; human; rat; osteoporosis;  
OS Bone deficiency; alveolar pyorrhea; fracture; vertebrate.  
FH Rattus rattus.  
FT Key  
FT Location/Qualifiers  
FT 1..366  
FT /label= sig\_peptide  
FT protein  
FT 367..476  
FT /label= mat\_protein  
FT /note= "claim 1, 3"  
FT peptide  
FT 1..365  
FT /label= sig\_peptide  
FT protein  
FT 366..476  
FT /label= mat\_protein  
FT peptide  
FT 1..364  
FT /label= sig\_peptide  
FT protein  
FT 365..476  
FT /label= mat\_protein  
FT PN W09401557-A.  
PD 20-JAN-1994.  
PF 09-JUL-1993; J00952.  
PR 13-JUL-1993; JP-206996.  
PA (SUMO) SUMITOMO MEXAL IND LTD. Keshita N, Konno Y;  
PI Fukuda K, Hino J, Kangawa K,  
PI Takao M;  
DR WPI; 94-035064/04.  
DR N-PSDB; Q54052.  
DE Bone formation-inducing protein - for therapy of diseases  
PT involving osteoporosis, a bone deficiency such as alveolar  
PT pyorrhea etc. and bone fracture  
PS Disclosure; Page 44-48; 57pp; English.  
CC Protein having improved bone formation inducing-activity  
CC has been provided. BIP mRNA may be obtained from the tissue  
CC of a vertebrate (e.g. human, rat) and used in recombinant  
CC DNA techniques for the prodn. of the protein. The BIP  
CC us useful in pharmaceuticals.  
SQ Sequence 476 AA;  
Query Match 75.0%; Score 42; DB 1; Length 476;  
Best Local Similarity 37.5%; Pred. No. 3.92e+02;  
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;  
Db 432 IVPGEIP 439  
QY 261 LIPSVDP 268  
RESULT 13  
ID R74970 standard; Protein; 476 AA.  
AC R74970;  
DE Murine growth differentiation factor-10 (GDF-10).  
KW Growth differentiation factor-10; transforming growth factor beta.  
OS Mus musculus.  
FH Key  
FT modified\_site 114  
FT /label= potential N-linked glycosylation site  
FT modified\_site 152  
FT /label= see above  
FT modified\_site 277  
FT /label= see above  
FT modified\_site 467  
FT /label= see above  
FT cleavage\_site 365  
FT /label= putative proteolytic processing site  
FT W09510539-A.  
PN 20-APR-1995.  
PD 07-OCT-1994; U11440.  
PF 08-OCT-1993; US-134078.  
PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MED.

PI Cunningham N, Lee S;  
DR WPI; 95-161740/21.  
DR N-PSDB; Q90457.  
DE New growth differentiation factor-10 - used to develop prods. for  
PT the detection and treatment of cell proliferative disorders.  
PS Example; Figure 2; 77pp; English.  
CC A murine uterus cDNA library consisting of 3x10(6) recombinant phage  
CC from the GDF-10 PCR product (see Q90454-Q90456). The entire nt  
CC sequence of the longest of 7 hybridising clones is shown in Q90457.  
CC It potentially encodes a protein with a mol. wt. of 52.5 kD. The  
CC predicted GDF-10 AA sequence contains a hydrophobic N-terminal  
CC region, suggestive of a signal sequence for secretion. Cleavage of  
CC the GDF-10 precursor would generate a mature protein with a predicted  
CC unglycosylated mol. wt. of 12.6 kD. The C-terminal region of GDF-10  
CC following the putative proteolytic processing site shows significant  
CC homology to the known members of the TGF-beta superfamily.  
SQ Sequence 476 AA;  
Query Match 75.0%; Score 42; DB 1; Length 476;  
Best Local Similarity 37.5%; Pred. No. 3.92e+02;  
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;  
Db 432 IVPGEIP 439  
QY 261 LIPSVDP 268  
RESULT 14  
ID W71468 standard; Protein; 550 AA.  
AC W71468;  
DE 11-JAN-1999 (first entry)  
DE Cercospora nicotianae cercosporin resistance crgl gene product.  
KW Photosensitiser resistance; cercosporin resistance; crgl gene;  
KW transgenic plant; crop protection; selectable marker.  
OS Cercospora nicotianae AFCC 18366.  
PN W09841082-A.  
PD 24-SEP-1998.  
PF 13-MAR-1998; U04981.  
PR 17-MAR-1997; US-040615.  
PA (UYNC-) UNIV NORTH CAROLINA STATE.  
PI Daub ME, Ehrenshaft M, Jenns AE;  
DR WPI; 98-520848/44.  
DR N-PSDB; V60573.  
DE New nucleic acid encoding protein that increases cellular resistance  
PT to photosensitiser - particularly cercosporin, used to protect  
PT plants against infection by Cercospora fungi, also related  
PT transformants and expression cassettes  
PS Claim 28; Page 39-41; 62pp; English.  
CC This novel protein is encoded by the crgl gene (see V60573) of  
CC Cercospora nicotianae, and upon expression provides or increases  
CC resistance to several photosensitisers, including cercosporin.  
CC Nucleic acid molecules (see also V60572) which, upon expression,  
CC increase resistance in a cell to a photosensitiser are claimed.  
CC they can be used to increase resistance of fungal or plant cells to  
CC photosensitisers (or to singlet oxygen species produced by  
CC photosensitisers), especially to improve resistance to  
CC photosensitiser-producing pathogens, particularly fungi of the  
CC genus Cercospora, specifically in plants (rape, canola, sorghum,  
CC soybean, sugarbeet, maize and tobacco), but also in bacteria, fungi  
CC and animals.  
SQ Sequence 550 AA;  
Query Match 75.0%; Score 42; DB 1; Length 550;  
Best Local Similarity 62.5%; Pred. No. 3.92e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Db 64 LFPVAVPAP 71  
QY 261 LIPSVDP 268  
RESULT 15

ID W48898 standard; Protein; 635 AA.  
 AC W48898.  
 DT 13-OCT-1998 (first entry)  
 DE Candida albicans Cabemip protein.  
 KW Cabemip; Bemip; rho family; screening; virulence;  
 KW hyphal formation; pathogenic fungi; inhibitor; inflammation;  
 KW antimycotic.  
 OS Candida albicans.  
 PN W09818927-A1.  
 PD 07-MAY-1998.  
 PF 29-OCT-1997; CA0809.  
 PR 30-OCT-1996; US-029458.  
 PA (CANA ) NAT RES COUNCIL CANADA.  
 PI Leberer E. Thomas DY;  
 DR WPI; 98-272222/24.  
 DR N-PSDB; V32556.  
 PT In vitro screening test for agents that inhibit Candida genes  
 PT involved in virulence - and transition to hyphal form, potentially  
 PT useful as antimycotic agents  
 PS Disclosure: Fig 12; 79pp; English.  
 CC The sequence is that of the Cabemip protein which can be used  
 CC in the development of an in vitro screening test for compounds  
 CC that inhibit biological activity of the protein and a system for  
 CC measuring its activity. The protein is involved in virulence and  
 CC hyphal formation. Inhibitors are potentially useful for rendering  
 CC pathogenic fungi (any species in which hyphal induction by kinase  
 CC occurs) avirulent and/or to treat inflammation.  
 SQ Sequence 635 AA;

Query Match 75.0%; Score 42; DB 1; Length 635;  
 Best Local Similarity 62.5%; Pred. No. 3.92e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 354 IIPSGP 361  
 QY 261 LIFSVPDP 268

Search completed: Wed May 10 12:09:37 2000  
 Job time : 8 secs.



```
SQ SEQUENCE 1170 AA; 129412 MW; 6503141 CN;
Query Match 80.4%; Score 45; DB 1; Length 1170;
Best Local Similarity 85.7%; Pred. No. 9.33e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 66 LIPVPDP 72
QY 261 LIPVPDP 267
|||||
RESULT 2
ID US-08-680-726A-12 STANDARD; PRT; 119 AA.
XX AC xxxxxx
XX DT
XX DE
XX SEQUENCE 12, Application US/08680726A
XX Sequence 12, Application US/08680726A
XX Patent No. 5804197
XX GENERAL INFORMATION:
XX APPLICANT: Haanes, Elizabeth J.
XX APPLICANT: Frank, Rexann S.
XX TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
XX NUMBER OF SEQUENCES: 92
XX CORRESPONDENCE ADDRESS:
XX ADDRESSEE: Sheridan Ross & McIntosh
XX STREET: 1700 Lincoln Street, Suite 3500
XX CITY: Denver
XX STATE: Colorado
XX COUNTRY: U.S.A.
XX ZIP: 80203
XX COMPUTER READABLE FORM:
XX MEDIUM TYPE: Floppy disk
XX COMPUTER: IBM PC compatible
XX OPERATING SYSTEM: PC-DOS/MS-DOS
XX SOFTWARE: Patent In Release #1.0, Version #1.30
XX CURRENT APPLICATION DATA:
XX APPLICATION NUMBER: US/08/680,726A
XX FILING DATE: 12-JUL-1996
XX CLASSIFICATION: 424
XX ATTORNEY/AGENT INFORMATION:
XX NAME: Connell, Gary J.
XX REGISTRATION NUMBER: 32,020
XX REFERENCE/DOCKET NUMBER: 2618-46-C1
XX TELECOMMUNICATION INFORMATION:
XX TELEPHONE: (303) 863-9700
XX TELEFAX: (303) 863-0223
XX INFORMATION FOR SEQ ID NO: 12:
XX SEQUENCE CHARACTERISTICS:
XX LENGTH: 119 amino acids
XX TYPE: amino acid
XX TOPOLOGY: linear
XX MOLECULE TYPE: protein
XX SEQUENCE 119 AA; 14041 MW; 78091 CN;
Query Match 76.8%; Score 43; DB 1; Length 119;
Best Local Similarity 50.0%; Pred. No. 1.53e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 33 MIPDIPNP 40
QY 261 LIPVPDP 268
|||||
RESULT 3
ID US-08-602-010A-12 STANDARD; PRT; 119 AA.
XX AC xxxxxx
XX DT
XX DE
XX SEQUENCE 19, Application US/08413118
XX Sequence 19, Application US/08413118
XX Patent No. 5685920
XX GENERAL INFORMATION:
XX APPLICANT: PAOLETTI, ENZO
XX APPLICANT: LIMBACH, KEITH J.
XX TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
XX CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
XX NUMBER OF SEQUENCES: 128
XX CORRESPONDENCE ADDRESS:
XX ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
XX STREET: 530 FIFTH AVENUE, 25TH FLOOR
XX CITY: NEW YORK
XX STATE: NEW YORK
XX COUNTRY: UNITED STATES OF AMERICA
```





CC LENGTH: 345 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC FRAGMENT TYPE: N-terminal  
SQ SEQUENCE 345 AA; 40609 MW; 660444 CN;

Query Match 76.8%; Score 43; DB 1; Length 345;  
Best Local Similarity 50.0%; Pred. No. 1.53e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 33 MIPDIPNP 40  
QY 261 LIPSVDPD 268

RESULT 7  
ID US-08-220-151-20 STANDARD; PRT; 345 AA.  
XX xxxxxx  
AC  
XX  
DT  
XX  
DE  
XX  
XX

Sequence 20, Application US/08220151  
Sequence 20, Application US/08220151  
Patent No. 5529780  
GENERAL INFORMATION:  
APPLICANT: Paolletti, Enzo  
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF  
TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC AND gD AND USES THEREFOR  
NUMBER OF SEQUENCES: 91  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/220,151  
FILING DATE: 30-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2540  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066 CURTMS  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 345 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
SEQUENCE 345 AA; 40609 MW; 660444 CN;

Query Match 76.8%; Score 43; DB 1; Length 345;  
Best Local Similarity 50.0%; Pred. No. 1.53e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 33 MIPDIPNP 40

QY 261 LIPSVDPD 268

RESULT 8  
ID US-08-220-151-19 STANDARD; PRT; 345 AA.  
XX xxxxxx  
AC  
XX  
DT  
XX  
DE  
XX  
XX

Sequence 19, Application US/08220151  
Sequence 19, Application US/08220151  
Patent No. 5529780  
GENERAL INFORMATION:  
APPLICANT: Paolletti, Enzo  
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF  
TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC AND gD AND USES THEREFOR  
NUMBER OF SEQUENCES: 91  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/220,151  
FILING DATE: 30-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2540  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066 CURTMS  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 345 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
SEQUENCE 345 AA; 40609 MW; 660444 CN;

Query Match 76.8%; Score 43; DB 1; Length 345;  
Best Local Similarity 50.0%; Pred. No. 1.53e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 33 MIPDIPNP 40  
QY 261 LIPSVDPD 268

RESULT 9  
ID US-08-680-726A-54 STANDARD; PRT; 345 AA.  
XX xxxxxx  
AC  
XX  
DT  
XX  
DE  
XX

Sequence 54, Application US/08680726A

```

CC Sequence 54, Application US/08680726A
CC Patent No. 5804197
CC GENERAL INFORMATION:
CC APPLICANT: Haanes, Elizabeth J.
CC APPLICANT: Frank, Rexann S.
CC TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES
CC NUMBER OF SEQUENCES: 92
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Sheridan Ross & McIntosh
CC STREET: 1700 Lincoln Street, Suite 3500
CC CITY: Denver
CC STATE: Colorado
CC COUNTRY: U.S.A.
CC ZIP: 80203
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION NUMBER: US/08/680,726A
CC FILING DATE: 12-JUL-1996
CC CLASSIFICATION: 424
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Connell, Gary J.
CC REGISTRATION NUMBER: 32,020
CC REFERENCE/DOCKET NUMBER: 2618-46-C1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (303) 863-9700
CC TELEFAX: (303) 863-0223
CC INFORMATION FOR SEQ ID NO: 54:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 345 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: Protein
CC SEQUENCE 345 AA: 40609 MW; 560444 CN;

Query Match 76.8%; Score 43; DB 1; Length 345;
Best Local Similarity 50.0%; Pred. No. 1.53e-02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 33 MIPDIPNP 40
QY 261 LIPSVDP 268

RESULT 10
ID US-08-497-599-19 STANDARD; PRT; 19 AA.
XX
AC xxxxxx
XX
DT
XX
DE
Sequence 19, Application US/08497599
Sequence 19, Application US/08497599
Patent No. 5908828
GENERAL INFORMATION:
APPLICANT: Kurita, Takashi
APPLICANT: Matsumoto, Tomoaki
APPLICANT: Kikuno, Reiko
APPLICANT: Otawara-Hamamoto, Yoko
APPLICANT: Breipohl, Gerhard
TITLE OF INVENTION: Synthetic Peptide Derivatives And The
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I. Street, N.W., Suite 700
CITY: Washington
STATE: D.C.

```

```

CC COUNTRY: USA
CC ZIP: 20005-3315
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION NUMBER: US/08/497,599
CC FILING DATE: 30-JUN-1995
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Toohey, Kimberlin M.
CC REGISTRATION NUMBER: 35,391
CC REFERENCE/DOCKET NUMBER: 02481.1451-00000
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202) 408-4000
CC TELEFAX: (202) 408-4400
CC INFORMATION FOR SEQ ID NO: 19:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 19 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: Peptide
CC SEQUENCE 19 AA: 1903 MW; 2338 CN;

Query Match 73.2%; Score 41; DB 2; Length 19;
Best Local Similarity 37.5%; Pred. No. 2.49e-02;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 12 VVVGPEP 19
QY 261 LIPSVDP 268

RESULT 11
ID US-08-335-583C-41 STANDARD; PRT; 103 AA.
XX
AC xxxxxx
XX
DT
XX
DE
Sequence 41, Application US/08335583C
Sequence 41, Application US/08335583C
Patent No. 5693779
GENERAL INFORMATION:
APPLICANT: Moos Jr., Malcolm
APPLICANT: Wang, Shouwan
APPLICANT: Krinks, Marie
TITLE OF INVENTION: PRODUCTION AND USE OF
TITLE OF INVENTION: ANTI-DORSALIZING MORPHOGENETIC PROTEIN
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq Version 1.5
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/335,583C
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:

```



US-09-376-430-2-08.ra

Thu May 11 06:49:41 2000

Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 60 VVPGIPEP 67  
: : : : :  
QY 261 LIPSVPDP 268

RESULT 14  
ID PCT-US93-08808-26 STANDARD; PRT: 104 AA.

XX  
AC  
XX  
XX  
DT  
XX

Sequence 26, Application PC/TUS9308808

Sequence 26, Application PC/TUS9308808  
GENERAL INFORMATION:

APPLICANT:  
TITLE OF INVENTION: MORPHOGEN-INDUCED LIVER REGENERATION  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CREATIVE BIOMOLECULES, INC.  
STREET: 45 SOUTH STREET  
CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 01748

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08808  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: KELLEY ESO, ROBIN D.  
REGISTRATION NUMBER: 34,637  
REFERENCE/DOCKET NUMBER: CRP-072  
TELEPHONE: 617/248-7477  
TELEFAX: 617/248-7100  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 104 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..104  
OTHER INFORMATION: /note= "BMP3"  
SEQUENCE 104 AA; 11552 MW; 61408 CN;

Query Match 73.2%; Score 41; DB 3; Length 104;  
Best Local Similarity 37.5%; Pred. No. 2.49e+02;  
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 60 VVPGIPEP 67  
: : : : :  
QY 261 LIPSVPDP 268

RESULT 15  
ID US-08-912-088-26 STANDARD; PRT: 104 AA.

XX  
AC  
XXXXXX

XX

DT

XX

DE

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Sequence 26, Application US/08912088  
Sequence 26, Application US/08912088  
Patent No. 5994131  
GENERAL INFORMATION:

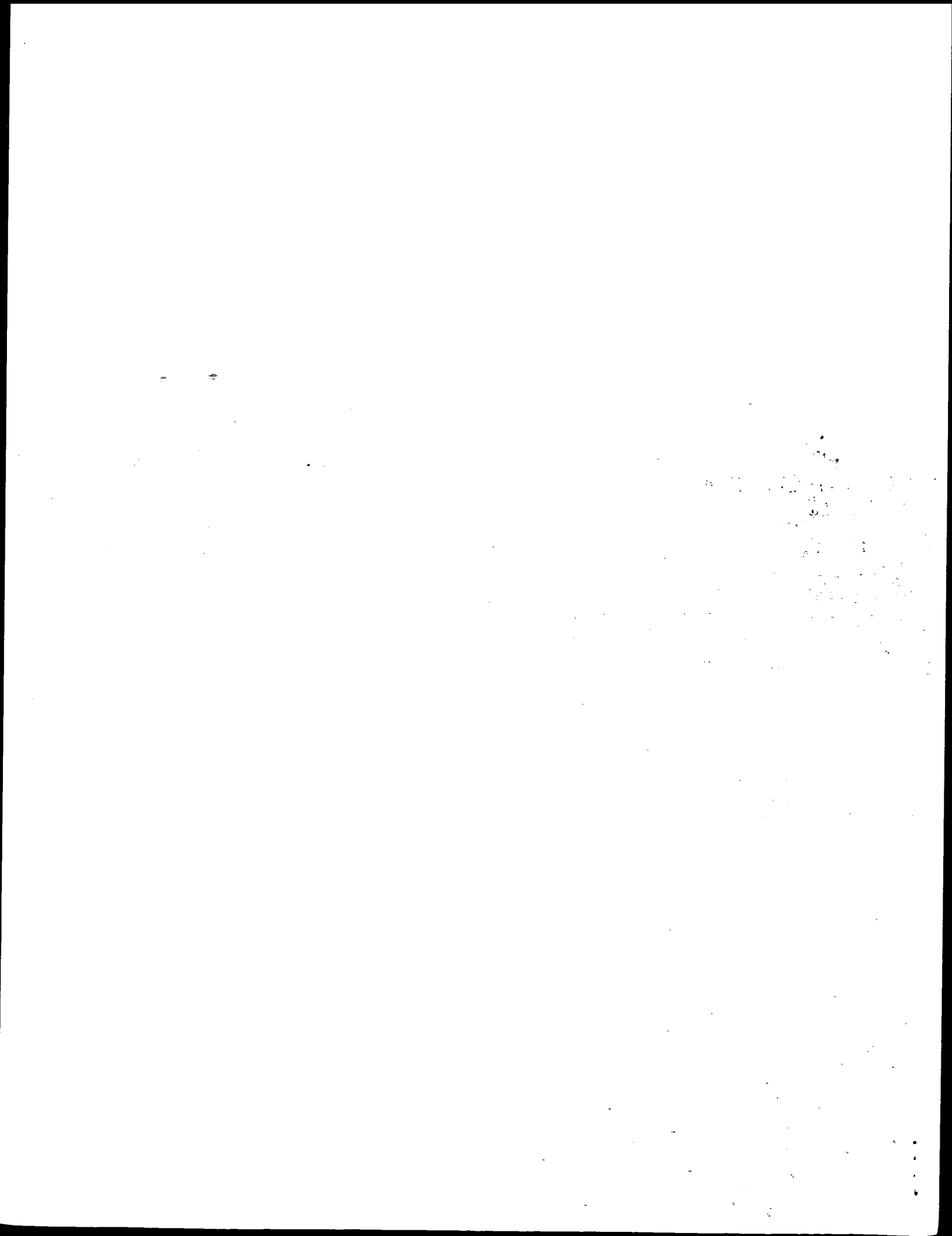
APPLICANT: SMART, JOHN  
APPLICANT: OPPERMAN, HERMAN  
APPLICANT: OZKAYNAK, ENGIN  
APPLICANT: KUBERASAMPATH, THANGAVEL  
APPLICANT: RUEGER, DAVID C.  
APPLICANT: PANG, ROY H.L.  
APPLICANT: COHEN, CHARLES M.  
TITLE OF INVENTION: MORPHOGENIC PROTEIN SCREENING METHOD  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
STREET: 45 SOUTH STREET  
CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 01748

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,088  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/278,729  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: PITCHER ESQ., EDMUND R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-058CPFW  
TELEPHONE: (508) 435-9001  
TELEFAX: (508) 435-6951  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 104 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: protein  
LOCATION: 1..104  
OTHER INFORMATION: /label= BMP3  
SEQUENCE 104 AA; 11552 MW; 61408 CN;

Query Match 73.2%; Score 41; DB 2; Length 104;  
Best Local Similarity 37.5%; Pred. No. 2.49e+02;  
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 60 VVPGIPEP 67  
: : : : :  
QY 261 LIPSVPDP 268

Search completed: Wed May 10 12:10:49 2000  
Job time : 55 sets,



\*\*\*\*\*  
 W P E S R L H  
 \*\*\*\*\*  
 (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 12:09:04 2000; MasPar time 3.92 Seconds  
 Tabular output not generated. 96.289 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (261-268) from US09376430A.pep (8 of 25)  
 Perfect score: 56  
 Sequence: 1 LIPSVDPDP 8

Scoring table: PAM 150  
 Gap 11

Searched: 142080 seqs, 47172405 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: p1r62

Statistics: 1:p1r1 2:p1r2 3:p1r3 4:p1r4

Mean 21.611; Variance 25.246; scale 0.856

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	ID	Description	Pred. No.
1	47	83.9	447	2 S66256	6.59e+00
2	46	82.1	258	2 T00311	1.08e+01
3	46	82.1	325	2 C73289	1.08e+01
4	46	82.1	721	2 A43275	1.08e+01
5	46	82.1	1170	2 A40558	1.08e+01
6	45	80.4	121	2 T11003	1.76e+01
7	45	80.4	229	2 S7957	1.76e+01
8	45	80.4	297	2 T00822	1.76e+01
9	45	80.4	320	2 A72452	1.76e+01
10	45	80.4	343	2 S34811	1.76e+01
11	45	80.4	343	2 S66173	1.76e+01
12	45	80.4	419	2 G70602	1.76e+01
13	45	80.4	662	2 T04856	1.76e+01
14	45	80.4	1170	1 TSHUP1	1.76e+01
15	45	80.4	3707	2 D18252	1.76e+01
16	44	78.6	142	2 D72695	2.84e+01
17	44	78.6	278	2 D70036	2.84e+01
18	44	78.6	391	2 T00823	2.84e+01
19	44	78.6	478	2 JC4838	2.84e+01
20	44	78.6	598	2 A55579	2.84e+01
21	44	78.6	600	2 F70980	2.84e+01
22	44	78.6	1515	2 B45502	2.84e+01
23	44	78.6	1767	2 A49502	2.84e+01

24 43 76.8 229 2 A69762 hypothetical protein 4.55e+01  
 25 43 76.8 334 2 C49348 succinoglycan biosynt 4.55e+01  
 26 43 76.8 348 2 S39958 exoO protein - Rhizob 4.55e+01  
 27 43 76.8 386 2 F72773 probable molybdopter 4.55e+01  
 28 43 76.8 443 2 S57328 uvsh protein - Emeric 4.55e+01  
 29 43 76.8 477 2 S55494 C3HC4 type zinc fing 4.55e+01  
 30 43 76.8 628 2 T02420 hypothetical protein 4.55e+01  
 31 43 76.8 646 2 I54546 immune-responsive gen 4.55e+01  
 32 43 76.8 664 2 T12988 hypothetical protein 4.55e+01  
 33 43 76.8 670 2 T02092 beta-fructofuranosid 4.55e+01  
 34 43 76.8 905 1 RGEVS5 regulatory protein SN 4.55e+01  
 35 43 76.8 910 2 S73361 DnaJ homolog protein 4.55e+01  
 36 43 76.8 1077 2 T01474 hypothetical protein 4.55e+01  
 37 43 76.8 1197 1 G65010 sensor protein evgs ( 4.55e+01  
 38 43 76.8 5825 2 T12117 polyprotein - favabe 4.55e+01  
 39 42 75.0 163 2 A14169 hypothetical protein 7.23e+01  
 40 42 75.0 187 2 F27000 probable dipeptide tr 7.23e+01  
 41 42 75.0 333 2 F26817 alpha-1,6-mannosyl-gl 7.23e+01  
 42 42 75.0 442 2 A57044 cytochrome b - Sulfol 7.23e+01  
 43 42 75.0 563 2 S21043 hypothetical protein 7.23e+01  
 44 42 75.0 640 2 S67656 hypothetical protein 7.23e+01  
 45 42 75.0 1054 2 T14189 hypothetical protein 7.23e+01

## ALIGNMENTS

RESULT 1

ENTRY S66256 #type complete  
 TITLE alpha-1,6-mannosyl-glycoprotein beta-1,  
 ORGANISM 2-N-acetylglucosaminyltransferase (EC 2.4.1.143) - human  
 DATE 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 07-May-1999 #

ACCESSIONS S66256  
 REFERENCE S66256  
 #authors Tan, J.; d'Agostaro, G.A.F.; Bendiak, B.; Reck, F.; Sarkar, M.; Squire, J.A.; Leong, P.; Schachter, H.  
 #journal Eur. J. Biochem. (1995) 231:317-328  
 #title The human UDP-N-acetylglucosamine:alpha-6-D-mannoside-beta-1,  
 2-N-acetylglucosaminyltransferase II gene (MGAT2). Cloning  
 of genomic DNA, localization to chromosome 14q21,  
 expression in insect cells and purification of the  
 recombinant protein.  
 #cross-references MIM:95361854  
 #accession S66256  
 ##molecule\_type DNA  
 ##residues 1-447 #label TAN  
 ##cross-references EMBL:U15128; NID:g902744; PID:g902745

GENETICS #gene GDB:MGAT2  
 #map\_position 14q21-14q21  
 KEYWORDS glycoprotein; glycosyltransferase; Golgi apparatus;  
 hexosyltransferase; transmembrane protein  
 FEATURE 10-28 #domain transmembrane #status predicted #label TMN  
 69,86 #binding-site carbohydrate (Asn) (covalent) #status predicted  
 SUMMARY #length 447 #molecular-weight 51550 #checksum 8407

Query Match 83.9%; Score 47; DB 2; Length 447;  
 Best Local Similarity 62.5%; Pred. No. 6.59e+00;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db '75 LIPAVPDP 82  
 QY 261 LIPSVDPDP 268

RESULT 2  
 ENTRY T00311 #type fragment  
 TITLE hypothetical protein, 31k - Escherichia coli plasmid p0157  
 Insertion sequence IS91(fragment)





```

REFERENCE      A42587      PID:g554390
#authors      Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin,
               M.F.; Dixit, V.M.
#journal      J. Biol. Chem. (1992) 267:3274-3281
#title        Characterization of mouse thrombospondin 2 sequence and
               expression during cell growth and development.
#cross-references MUID:92147683
#accession    B42587
##status      preliminary; not compared with conceptual translation
##molecule_type mRNA
##residues     1-1152/'P',1154-1170 #label LAH
##cross-references GB:M87276
##note         sequence extracted from NCBI backbone (NCBIP:81501)

REFERENCE      S6878
#authors      Chen, H.; Aeschlimann, D.; Nowlen, J.; Mosher, D.F.
#journal      FES Lett. (1996) 387:36-41
#title        Expression and initial characterization of recombinant mouse
               thrombospondin 1 and thrombospondin 3.
#cross-references MUID:96234006
#accession    S6878
##molecule_type protein
##residues     19-26/'X',28-37 #label CHE
COMPLEX        homotrimer; disulfide linked
CLASSIFICATION #superfamily thrombospondin 1; EGF homology; thrombospondin
               type 1 repeat homology; von Willebrand factor type C repeat
               homology
KEYWORDS        calcium binding; glycoprotein; homotrimer
FEATURE
1-18            #domain signal sequence #status predicted #label SIG\
19-1170         #product thrombospondin 1 #status predicted #label MAT\
317-375         #domain von Willebrand factor type C repeat homology
               #label VWC\
378-429         #domain thrombospondin type 1 repeat homology #label
               THR1\
434-490         #domain thrombospondin type 1 repeat homology #label
               THR2\
491-547         #domain thrombospondin type 1 repeat homology #label
               THR3\
551-586         #domain EGF homology #label EGF\
248,360,708,1067 #binding-site carbohydrate (Asn) (covalent) #status
               predicted
SUMMARY         #length 1170 #molecular-weight 129646 #checksum 3974

Query Match      82.1%; Score 46; DB 2; Length 1170;
Best Local Similarity 85.7%; Pred. No. 1.08e+01;
Matches          6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db      66 LIPAVPD 72
QY      261 LIPSVDP 267

RESULT      6
ENTRY       T11003      #type complete
TITLE       MLC1536.11 protein - Mycobacterium leprae
ORGANISM    #formal_name Mycobacterium leprae
DATE        16-Jul-1998 #sequence_revision 16-Jul-1999 #text_change
               16-Jul-1999
ACCESSIONS  T11003
REFERENCE    216918
#authors     Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
#submission  submitted to the EMBL Data Library, September 1997
#accession   T11003
##status     preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues   1-121 #label PAR
##cross-references EMBL:Z99125; NID:el061170; PID:e343536;
               PIDN:CAB16154.1

GENETICS
#gene        MLC1536.11
SUMMARY      #length 121 #molecular-weight 13504 #checksum 7722

```

```

Query Match      80.4%; Score 45; DB 2; Length 121;
Best Local Similarity 62.5%; Pred. No. 1.76e+01;
Matches          5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db      61 VIPGVPEP 68
QY      261 LIPSVDP 268

RESULT      7
ENTRY       S57957      #type fragment
TITLE       thrombospondin 1 - bovine (fragment)
ORGANISM    #formal_name Bos primigenius taurus #common_name cattle
DATE        13-Jan-1996 #sequence_revision 19-Apr-1996 #text_change
               20-Aug-1999
ACCESSIONS  S57957
REFERENCE    S57955
#authors     Lafuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.;
               Feige, J.J.
#submission  submitted to the EMBL Data Library, July 1995
#description Opposite regulation of thrombospondin-1 and
               CRISP/thrombospondin-2 expression by ACTH in adrenocortical
               cells.
#accession   S57957
##status     preliminary
##molecule_type mRNA
##residues   1-229 #label LAF
##cross-references EMBL:X89511; NID:g899228; PIDN:CAA61682.1;
               PID:g899229
CLASSIFICATION #superfamily thrombospondin 1; EGF homology; thrombospondin
               type 1 repeat homology; von Willebrand factor type C repeat
               homology
SUMMARY      #length 229 #checksum 5431

Query Match      80.4%; Score 45; DB 2; Length 229;
Best Local Similarity 85.7%; Pred. No. 1.76e+01;
Matches          6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db      48 LIPVDP 54
QY      261 LIPSVDP 267

RESULT      8
ENTRY       T00822      #type complete
TITLE       hypothetical protein T32G6.18 - Arabidopsis thaliana
ORGANISM    #formal_name Arabidopsis thaliana #common_name mouse-ear
               cress
DATE        12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change
               14-May-1999
ACCESSIONS  T00822
REFERENCE    Z14159
#authors     Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.;
               Brandon, R.C.; Sykes, S.M.; Kaul, S.; Mason, T.M.;
               Kerlavage, A.R.; Adams, M.D.; Somerville, C.R.; Venter,
               J.C.
#submission  submitted to the EMBL Data Library, November 1997
#description Arabidopsis thaliana chromosome II BAC T32G6 genomic
               sequence.
#accession   T00822
##status     translated from GB/EMBL/DBJ
##molecule_type DNA
##residues_type 1-297 #label ROU
##cross-references EMBL:AC002510; NID:g2618683; PID:g2618701
               #experimental_source cultivar Columbia
GENETICS
#map_position 2
#note         T32G6.18
SUMMARY      #length 297 #molecular-weight 32510 #checksum 6984

Query Match      80.4%; Score 45; DB 2; Length 297;
Best Local Similarity 71.4%; Pred. No. 1.76e+01;
Matches          5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 40 LIPTYPE 46
    |||||:
QY 261 LIPSVDP 267

RESULT 9
ENTRY #type complete
TITLE Probable peptide transport ATP-binding protein APE2261 -
ORGANISM Aeropyrum pernix (strain K1)
DATE #formal_name Aeropyrum pernix
20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change
A72452
A72450
A72452
Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.;
Haikawa, Y.; Jin-no, K.; Takahashi, M.; Sekine, M.; Baba,
S.; Ankaei, A.; Kosugi, H.; Hosoyama, A.; Fukui, S.; Nagai,
Y.; Nishijima, K.; Nakazawa, H.; Takamiya, M.; Masuda, S.;
Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.;
Kushida, N.; Oguchi, A.; Aoki, K.; Kubota, K.; Nakamura,
Y.; Nomura, N.; Sako, Y.; Kikuchi, H.
#journal DNA Res. (1999) 6:83-101
#title Complete genome sequence of an aerobic hyper-thermophilic
#cross-references MUID:99310339
#accession A72452
#status preliminary
#molecule_type DNA
#residues 1-320 #label RAW
#cross-references DBJ:AP000064; NID:g5105945; PIDN:BAA81273.1;
#experimental_source strain K1

GENETICS
#gene APE2261
#length 320 #molecular-weight 34740 #checksum 2031
SUMMARY
Query Match 80.4%; Score 45; DB 2; Length 320;
Best Local Similarity 87.5%; Pred. No. 1.76e+01;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 256 LIKSVDP 263
    |||||:
QY 261 LIPSVDP 268

RESULT 10
ENTRY #type complete
TITLE mating factor MAT1-2 - fungus (Cochliobolus heterostrophus)
ORGANISM #formal_name Cochliobolus heterostrophus, Bipolaris maydis
DATE 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
24-Sep-1999
A72452
A72450
A72452
Turgeon, B.G.; Bohlmann, H.; Ciuffetti, L.M.; Christiansen,
S.K.; Yang, G.; Schaefer, W.; Yoder, O.C.
#journal Mol. Gen. Genet. (1993) 238:270-284
#title Cloning and analysis of the mating type genes from
#cross-references MUID:93241164
#accession S34811
#status preliminary
#molecule_type DNA
#residues 1-343 #label TUR
#cross-references EMBL:X68398; NID:g2505; PIDN:CAA48464.1; PID:g2586

GENETICS
#gene MAT1-2
#introns 162/1
#classification #superfamily unassigned HMG box proteins; HMG box homology
#keywords DNA binding; nucleus; transcription regulation
#feature 129-203
#domain HMG box homology #label HMG1
#length 343 #molecular-weight 38319 #checksum 1452
SUMMARY

```

```

Query Match 80.4%; Score 45; DB 2; Length 343;
Best Local Similarity 71.4%; Pred. No. 1.76e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 237 LIPTYPE 243
    |||||:
QY 261 LIPSVDP 267

RESULT 11
ENTRY #type complete
TITLE mating factor MAT-2 homolog - Bipolaris sacchari
ALTERNATE_NAMES DNA-binding protein MAT-2
ORGANISM #formal_name Bipolaris sacchari
DATE 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change
24-Sep-1999
S66173
S66173
#accessions S66173
#authors Sharon, A.; Yamaguchi, K.; Christiansen, S.; Horwitz, B.A.;
Yoder, O.C.; Turgeon, B.G.
#journal Mol. Gen. Genet. (1996) 251:60-68
#title An asexual fungus has the potential for sexual development.
#cross-references MUID:96213991
#accession S66173
#status nucleic acid sequence not shown
#molecule_type DNA
#residues 1-343 #label SHA
#cross-references EMBL:X95814; NID:g1212792; PIDN:CAA65081.1;
#experimental_source strain 764-1

GENETICS
#gene MAT-2
#introns 162/1
#classification #superfamily unassigned HMG box proteins; HMG box homology
#keywords DNA binding
#feature 128-203
#domain HMG box homology #label HMG1
#length 343 #molecular-weight 38257 #checksum 782
SUMMARY
Query Match 80.4%; Score 45; DB 2; Length 343;
Best Local Similarity 71.4%; Pred. No. 1.76e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 237 LIPTYPE 243
    |||||:
QY 261 LIPSVDP 267

RESULT 12
ENTRY #type complete
TITLE hypothetical protein RV1004c - Mycobacterium tuberculosis
ORGANISM #formal_name Mycobacterium tuberculosis
DATE 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change
17-Jul-1998
G70602
A70500
#accessions G70602
#authors Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher,
C.; Harris, D.; Gordon, S.V.; Eigmeier, K.; Gas, S.; Barry
III, C.E.; Tekala, F.; Badcock, K.; Basham, D.; Brown, D.;
Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;
Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.;
Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.;
Skellern, S.; Squares, S.; Squires, R.; Sulston, J.E.;
Taylor, K.; Whitehead, S.; Barrall, B.G.
#journal Nature (1998) 393:537-544
#title Deciphering the biology of Mycobacterium tuberculosis from
#cross-references MUID:98295987
#accession G70602

```

```
##status      preliminary; nucleic acid sequence not shown;
               translation not shown
##molecule_type DNA
##residues    1-419 #label COL
##cross-references GB:294752; GB:AL123456; NID:g3261731; PID:e1299781;
               PID:g3261734
##experimental_source strain H37Rv

GENETICS
#gene
#summary
  Rv1004c
  #length 419 #molecular_weight 38785 #checksum 4947
  Query Match      80.4%; Score 45; DB 2; Length 419;
  Best Local Similarity 85.7%; Pred. No. 1.76e+01;
  Matches          6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 161 IPGVDPDP 167
   ||:||||
QY 262 IPSVPDP 268

RESULT 13
ENTRY T04856 #type fragment
TITLE hypothetical protein F28A21.50 - Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
          cress
DATE 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change
      11-Jun-1999
ACCESSIONS T04856
REFERENCE 215387
#authors Bevan, M.; Mueller, M.W.; Muendlein, A.; Felber, R.;
          Bancroft, I.; Wewes, H.W.; Mayer, K.F.X.; Schueller, C.
#submission submitted to the Protein Sequence Database, February 1999
#accession T04856
#molecule_type DNA
#residues 1-662 #label BEV
##cross-references EMBL:AL035526
##experimental_source cultivar Columbia; BAC clone F28A21

GENETICS
#map_position 4
#note Intron positions not resolved
#note F28A21.50
#summary #length 662 #checksum 4645

Query Match      80.4%; Score 45; DB 2; Length 662;
Best Local Similarity 62.5%; Pred. No. 1.76e+01;
Matches          5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 263 LVPPVPSP 270
   |::|||
QY 261 LIPVPDP 268

RESULT 14
ENTRY TSHUPL
TITLE thrombospondin 1 precursor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change
      13-Aug-1999
ACCESSIONS A26155; A34274; A30140; A25812; A05172; A42927
REFERENCE A26155
#authors Lawler, J.; Hynes, R.O.
#journal J. Cell Biol. (1986) 103:1635-1648
#title The structure of human thrombospondin, an adhesive
          glycoprotein with multiple calcium-binding sites and
          homologies with several different proteins.
#cross-references MUID:87057617
#accession A26155
#molecule_type mRNA
#residues 1-1170 #label LAW
##cross-references GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138
##note parts of this sequence, including the amino end of the
          mature protein, were determined by protein sequencing
```

```
REFERENCE A34274
#authors Laherty, C.D.; Gierman, T.M.; Dixit, V.M.
#journal J. Biol. Chem. (1989) 264:11222-11227
#title Characterization of the promoter region of the human
          thrombospondin gene. DNA sequences within the first intron
          increase transcription.
#cross-references MUID:89291870
#accession A34274
#molecule_type DNA
#residues 1-166 #label LAH
##cross-references GB:J04835
REFERENCE A30140
#authors Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.;
          Baumgarcel, D.M.; Rotwein, P.; Frazier, W.A.
#journal J. Cell Biol. (1989) 108:729-736
#title Complete thrombospondin mRNA sequence includes potential
          regulatory sites in the 3' untranslated region.
#cross-references MUID:89139590
#accession A30140
#molecule_type mRNA
#residues 1-83; 'A', 85-522; 'A', 524-1170 #label HEN
##cross-references EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465
#note parts of this sequence, including the amino end of the
          mature protein, were determined by protein sequencing
REFERENCE A25812
#authors Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.
#journal Biochemistry (1986) 25:8418-8425
#title Partial amino acid sequence of human thrombospondin as
          determined by analysis of cDNA clones: homology to malarial
          circumsporozoite proteins.
#cross-references MUID:87157592
#accession A25812
#molecule_type mRNA
#residues 1-83; 'A', 85-397 #label KOB
##cross-references GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354
REFERENCE A05172
#authors Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.;
          Frazier, W.A.
#journal Proc. Natl. Acad. Sci. U.S.A. (1986) 83:5449-5453
#cross-references MUID:86287276
#accession A05172
#molecule_type mRNA
#residues 1-83; 'A', 85-374; 'RC' #label DIX
##cross-references GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801
#note parts of this sequence, including the amino end of the
          mature protein, were determined by protein sequencing
REFERENCE A42927
#authors Sun, X.; Skorstengaard, K.; Mosher, D.F.
#journal J. Cell Biol. (1992) 118:693-701
#title Disulfides modulate RGD-inhibitable cell adhesive activity of
          thrombospondin
#cross-references MUID:92348511
#accession A42927
#molecule_type protein
#residues 987-1003 #label SUN
##note Cys-992 is shown to have a free sulfhydryl
GENETICS
#gene GDB:THBS1; TSP1; TSP
#map_position 15q15-15q15
#introns 23/1
#note the list of introns may be incomplete
#complex homotrimer, disulfide linked
#function participates in cell migration and adhesion, and in platelet
          aggregation
#description #superfamily thrombospondin 1; EGF homology; thrombospondin
          type 1 repeat homology; von Willebrand factor type C repeat
          homology
CLASSIFICATION
#keywords beta-hydroxyasparagine; calcium binding; cell adhesion;
          glycoprotein; trimer
#feature
#note #domain signal sequence #status predicted #label SIG\
      1-18
```

```

19-1170 #product thrombospondin 1 #status predicted #label MAT\
317-375 #domain von Willebrand factor type C repeat homology
#label VWC\
378-429 #domain thrombospondin type 1 repeat homology #label
THR1\
434-490 #domain thrombospondin type 1 repeat homology #label
THR2\
491-547 #domain thrombospondin type 1 repeat homology #label
THR3\
551-586 #domain EGF homology #label EGF\
650-689 #domain EGF homology #label EGF\
926-928 #region cell attachment (R-G-D) motif\
171-232 #disulfide_bonds #status predicted\
248,360,708,1067 #binding_site carbohydrate (Asn) (covalent) #status
predicted\
270,274 #disulfide_bonds interchain #status predicted\
610 #modified_site erythro-beta-hydroxyasparagine (Asn)
#status predicted\
1051 #binding_site carbohydrate (Asn) (covalent) #status
absent
SUMMARY #length 1170 #molecular-weight 129412 #checksum 6656
Query Match 80.4%; Score 45; DB 1; Length 1170;
Best Local Similarity 85.7%; Pred. No. 1.76e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 66 LIPSPVPD 72
QY 261 LIPSPVPD 267

```

```

RESULT 15
ENTRY S18252 #type complete
TITLE heparan sulfate proteoglycan - mouse
ALTERNATE_NAMES perlecan
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
13-Aug-1999
ACCESSIONS S18252; A31917; B31917; S66460
REFERENCE S18252
#authors Noonan, D.M.; Fulle, A.; Valente, P.; Cai, S.; Horigan, E.;
Sasaki, M.; Yamada, Y.; Hassell, J.R.
#journal J. Biol. Chem. (1991) 266:22939-22947
#title The complete sequence of perlecan, a basement membrane
heparan sulfate proteoglycan, reveals extensive similarity
with laminin A chain, low density lipoprotein-receptor, and
the neural cell adhesion molecule.
#cross-references MUID:92078153
#accession S18252
#molecule_type mRNA
#residues 1-3707 ##label NOO
##cross-references EMBL:M77174; NID:g200295; PID:g200296
REFERENCE A92680
#authors Noonan, D.M.; Horigan, E.A.; Ledbetter, S.R.; Vogeli, G.;
Sasaki, M.; Yamada, Y.; Hassell, J.R.
#journal J. Biol. Chem. (1988) 263:16379-16387
#title Identification of cDNA clones encoding different domains of
the basement membrane heparan sulfate proteoglycan.
#cross-references MUID:89034110
#accession A31917
#molecule_type mRNA
#residues 940-1601 ##label NO2
##cross-references GB:J04054; NID:g200252; PID:g200253
#accession B31917
#molecule_type mRNA
#residues 1870-2600 ##label NO3
##cross-references GB:J04055; NID:g200300; PID:g200301
#accession S66460
#authors Schulze, B.; Mann, K.; Battistutta, R.; Wiedemann, H.; Timpl,
R.
#journal Eur. J. Biochem. (1995) 231:551-556
#title Structural properties of recombinant domain III-3 of perlecan
containing a globular domain inserted into an

```

```

#cross-references epidermal-growth-factor-like motif.
#accession S66460
#molecule_type protein
#residues 1272-1274, 'X', 1276, 'X', 1278-1279 ##label SCH
CLASSIFICATION #superfamily LDL receptor ligand-binding repeat homology; EGF
homology; laminin G repeat homology; laminin-type EGF-like
homology
KEYWORDS glycoprotein
FEATURE
199-234 #domain LDL receptor ligand-binding repeat homology
#label LDL1\
285-319 #domain LDL receptor ligand-binding repeat homology
#label LDL2\
325-359 #domain LDL receptor ligand-binding repeat homology
#label LDL3\
368-403 #domain LDL receptor ligand-binding repeat homology
#label LDL4\
764-811 #domain laminin-type EGF-like homology #label LEG\
1159-1206 #domain laminin-type EGF-like homology #label LEG7\
1563-1610 #domain laminin-type EGF-like homology #label EG7\
1613-1668 #domain laminin-type EGF-like homology #label LEG8\
3163-3198 #domain EGF homology #label EGF\
3270-3423 #domain laminin G repeat homology #label LG2\
3464-3492 #domain EGF homology #label EGF7\
1256,1891,2336,
2394,2427
#binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY #length 3707 #molecular-weight 398291 #checksum 1636

```

```

Query Match 80.4%; Score 45; DB 2; Length 3707;
Best Local Similarity 75.0%; Pred. No. 1.76e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

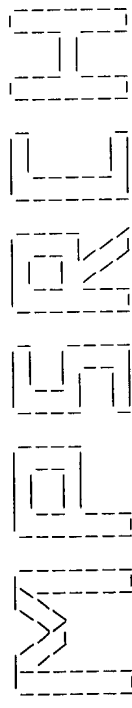
Db 275 LIPSPVPD 282
QY 261 LIPSPVPD 268

```

```

Search completed: Wed May 10 12:09:12 2000
Job time : 8 secs.

```



Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 12:02:51 2000; MasPar time 92.54 Seconds  
Tabular output not generated. 2.633 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (261-268) from US09376430A.pap (8 of 25)  
Perfect Score: 56  
Sequence: 1 LIPSVDP 8

Scoring table: PAM 150  
Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot38  
1:swissprot

Statistics: Mean 22.121; Variance 22.225; scale 0.995

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	47	83.9	447	1	GNT2_HUMAN	1.67e+00
2	46	82.1	1170	1	TSPI_MOUSE	2.92e+00
3	45	80.4	343	1	MAT2_COCH	5.07e+00
4	45	80.4	525	1	V12_HPV60	5.07e+00
5	45	80.4	1073	1	V126_HUMAN	5.07e+00
6	45	80.4	1170	1	TSPI_HUMAN	5.07e+00
7	45	80.4	3707	1	PGEM_MOUSE	5.07e+00
8	44	78.6	478	1	BM3B_HUMAN	8.70e+00
9	44	78.6	598	1	BCCA_MYCLE	8.70e+00
10	43	76.8	334	1	EXOO_RHIME	1.48e+01
11	43	76.8	646	1	IRGL_MOUSE	1.48e+01
12	43	76.8	670	1	INV1_MAIZE	1.48e+01
13	43	76.8	905	1	SNF5_YEAST	1.48e+01
14	43	76.8	910	1	DNM1_MYCPN	1.48e+01
15	43	76.8	1197	1	EVGS_ECOLI	1.48e+01
16	42	75.0	181	1	ATP SYNTHASE B CHAIN	2.49e+01
17	42	75.0	305	1	MALM_SALTY	2.49e+01
18	42	75.0	306	1	MALM_ECOLI	2.49e+01
19	42	75.0	333	1	YAKU_RHISN	2.49e+01
20	42	75.0	442	1	PROBABLE ARC TRANSPORT	2.49e+01
21	42	75.0	462	1	ERG_HUMAN	2.49e+01
22	42	75.0	476	1	TRANSCRIPTIONAL REGULA	2.49e+01
23	42	75.0	476	1	BONE MORPHOGENETIC PRO	2.49e+01
					BONE MORPHOGENETIC PRO	2.49e+01

RESULT ID	GNT2_HUMAN	STANDARD	PRT	447 AA
AC	Q10469			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	ALPHA-1,6-MANNOSYL-GLYCOPROTEIN BETA-1,2-N-			
DE	ACETYLGLUCOSAMINYLTRANSFERASE (EC 2.4.1.143) (N-GLYCOSYL-			
DE	OLIGOSACCHARIDE-GLYCOPROTEIN N-ACETYLGLUCOSAMINYLTRANSFERASE II)			
DE	(BETA-1,2-N-ACETYLGLUCOSAMINYLTRANSFERASE II) (GNT-II) (GLCNAC-T II).			
GN	MGAT2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
OC	Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-16.			
RC	TISSUE=LEUKOCYTE;			
RA	MEDLINE; 95361854.			
RA	Tan J., D'Agostaro A.F., Bendiak B., Reck F., Sarkar M., Squire J.A.,			
RA	Leong P., Schachter H.;			
RT	"The human UDP-N-acetylglucosamine: alpha-6-D-mannoside-beta-1,2-N-			
RT	acetylglucosaminyltransferase II gene (MGAT2). Cloning of genomic			
RT	DNA, localization to chromosome 14q21, expression in insect cells and			
RT	purification of the recombinant protein."			
RL	Eur. J. Biochem. 231:317-328(1995).			
RN	[2]			
RP	VARIANTS CDG2 ARG-262 AND PHE-290.			
RA	MEDLINE; 96404413.			
RA	Tan J., Dunn J., Jaeken J., Schachter H.;			
RT	"Mutations in the MGAT2 gene controlling complex N-glycan synthesis			
RT	cause carbohydrate-deficient glycoprotein syndrome type II, an			
RT	autosomal recessive disease with defective brain development."			
RL	Am. J. Hum. Genet. 59:810-817(1996).			
CC	-1- FUNCTION: CATALYZES AN ESSENTIAL STEP IN THE CONVERSION OF OLIGO-			
CC	MANNOSE TO COMPLEX N-GLYCANS.			
CC	-1- CATALYTIC ACTIVITY: UDP-N-ACETYL-D-GLUCOSAMINE + ALPHA-D-MANNOSYL-			
CC	1,6-(N-ACETYL-BETA-D-GLUCOSAMINYL-1,2-ALPHA-D-MANNOSYL-1,3)-BETA-			
CC	D-MANNOSYL-R = UDP + N-ACETYL-BETA-D-GLUCOSAMINYL-1,2-ALPHA-D-			
CC	MANNOYL-1,6-(N-ACETYL-BETA-D-GLUCOSAMINYL-1,2-ALPHA-D-MANNOSYL-			
CC	1,3)-BETA-D-MANNOSYL-R.			
CC	-1- PATHWAY: GLYCOSYLATION.			
CC	-1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI.			
CC	-1- DISEASE: CONGENITAL DEFICIENCY OF MGAT2 IS ASSOCIATED WITH SEVERE			
CC	IMPAIRMENT OF NORMAL EMBRYOGENESIS, PARTICULARLY IN THE NERVOUS			
CC	SYSTEM IN PATIENTS WITH CARBOHYDRATE-DEFICIENT GLYCOPROTEIN			
CC	SYNDROME II (CDGS TYPE II OR CDG2). IT HAS ALSO BEEN IMPLICATED IN			
CC	THE ETIOLOGY OF A VARIANT FORM OF CONGENITAL DYSPERYTHROPOIETIC			

24	42	75.0	563	1	CYB_SULAC	2.49e+01
25	42	75.0	2150	1	SDC3_CAEEL	2.49e+01
26	41	73.2	80	1	BMP3_BOVIN	4.15e+01
27	41	73.2	362	1	MURG_STRCU	4.15e+01
28	41	73.2	364	1	MURG_STRCO	4.15e+01
29	41	73.2	415	1	ILSR_MOUSE	4.15e+01
30	41	73.2	429	1	GAG_HTL1A	4.15e+01
31	41	73.2	429	1	GAG_HTL1M	4.15e+01
32	41	73.2	429	1	GAG_HTL1C	4.15e+01
33	41	73.2	456	1	YRDX_SHOSH	4.15e+01
34	41	73.2	476	1	BMP3_HUMAN	4.15e+01
35	41	73.2	476	1	LEU2_MYCLE	4.15e+01
36	41	73.2	490	1	IRF5_MOUSE	4.15e+01
37	41	73.2	497	1	IRF5_MOUSE	4.15e+01
38	41	73.2	527	1	TKX_MOUSE	4.15e+01
39	41	73.2	529	1	CIK6_HUMAN	4.15e+01
40	41	73.2	530	1	CIK6_RAI	4.15e+01
41	41	73.2	703	1	COAT_SMSV4	4.15e+01
42	41	73.2	836	1	GCSR_HUMAN	4.15e+01
43	41	73.2	1024	1	UBAL_YEAST	4.15e+01
44	41	73.2	1536	1	SIN3_YEAST	4.15e+01
45	40	71.4	432	1	PYRP_BACCL	6.83e+01

ALIGNMENTS

ANEMIA TYPE II.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; U15128; AAA86956.1; -  
 DR MIN; 602616; -  
 DR MIN; 212066; -  
 KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;  
 KW Glycoprotein; Golgi stack; Disease mutation.  
 FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 10 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).  
 FT DOMAIN 30 447 LUMENAL, CATALYTIC (POTENTIAL).  
 FT CARBOHYD 69 69 POTENTIAL.  
 FT CARBOHYD 86 86 H -> R (IN CDG2).  
 FT VARIANT 262 262 S -> F (IN CDG2).  
 FT VARIANT 290 290 /FtId=VAR\_003415.  
 FT VARIANT 290 290 /FtId=VAR\_003416.  
 SQ SEQUENCE 447 AA; 51550 MW; 533B76D08B8A572 CRC64;  
 Query Match 83.9%; Score 47; DB 1; Length 447;  
 Best Local Similarity 62.5%; Pred. No. 1.67e+00;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 75 LVPVPOQ 82  
 | : | : |  
 QY 261 LIPSVPDP 268

RESULT 2  
 ID TSPL\_MOUSE STANDARD; PRT; 1170 AA.  
 AC P35441;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE THROMBOSPONDIN 1 PRECURSOR.  
 GN THBS1 OR TSPL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92128941.  
 RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,  
 RA Jenkins N.A.;  
 RL "Characterization of the murine thrombospondin gene.";  
 RL Genomics 11:587-600(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92147683.  
 RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,  
 RA Dixit V.M.;  
 RL "Characterization of mouse thrombospondin 2 sequence and expression during cell growth and development.";  
 RL J. Biol. Chem. 267:3274-3281(1992).  
 RN [3]  
 RP SEQUENCE OF 1-490 FROM N.A.  
 RX MEDLINE; 90375546.  
 RA Bornstein P., Alfi D., Devaravala S., Framson P., Li P.;  
 RL "Characterization of the mouse thrombospondin gene and evaluation of the role of the first intron in human gene expression.";  
 RL J. Biol. Chem. 265:16691-16698(1990).  
 CC -!- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND CELL-TO-MATRIX INTERACTIONS. CAN BIND TO FIBRINOGEN, FIBRONECTIN, LAMININ AND TYPE V COLLAGEN.  
 CC -!- SUBUNIT: HOMOTRIMER, CROSS-LINKED BY DISULFIDE BONDS.  
 CC -!- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.

CC -!- SIMILARITY: CONTAINS 1 VMFC DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 3 TYPE-1 TSP REPEATS AND 7 TYPE-3 TSP REPEATS (WHICH BIND CALCIUM).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; M62470; AAA50611.1; -  
 DR EMBL; M62450; AAA50611.1; JOINED.  
 DR EMBL; M62451; AAA50611.1; JOINED.  
 DR EMBL; M62452; AAA50611.1; JOINED.  
 DR EMBL; M62453; AAA50611.1; JOINED.  
 DR EMBL; M62454; AAA50611.1; JOINED.  
 DR EMBL; M62455; AAA50611.1; JOINED.  
 DR EMBL; M62456; AAA50611.1; JOINED.  
 DR EMBL; M62457; AAA50611.1; JOINED.  
 DR EMBL; M62458; AAA50611.1; JOINED.  
 DR EMBL; M62459; AAA50611.1; JOINED.  
 DR EMBL; M62460; AAA50611.1; JOINED.  
 DR EMBL; M62461; AAA50611.1; JOINED.  
 DR EMBL; M62462; AAA50611.1; JOINED.  
 DR EMBL; M62463; AAA50611.1; JOINED.  
 DR EMBL; M62464; AAA50611.1; JOINED.  
 DR EMBL; M62465; AAA50611.1; JOINED.  
 DR EMBL; M62466; AAA50611.1; JOINED.  
 DR EMBL; M62467; AAA50611.1; JOINED.  
 DR EMBL; M62468; AAA50611.1; JOINED.  
 DR EMBL; M62469; AAA50611.1; JOINED.  
 DR EMBL; M62476; AAA50611.1; JOINED.  
 DR EMBL; J05606; AAA40431.1; -  
 DR EMBL; J05605; AAA40431.1; -  
 DR PIR; A40558; A40558.  
 DR PIR; B42587; B42587.  
 DR PIR; A37905; A37905.  
 DR HSP; P35555; LEMO.  
 DR MGD; MGI:98737; THBS1.  
 DR PROSITE; PS00222; EGF\_1; FALSE\_NEG.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01208; VMFC; 1.  
 DR PFAM; PF00008; EGF; 1.  
 DR PFAM; PF00090; tsp\_1; 3.  
 DR PFAM; PF00093; vwc; 1.  
 KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;  
 KW EGF-like domain; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 1170  
 FT DOMAIN 19 232 THROMBOSPONDIN 1.  
 FT DOMAIN 316 373 HEPARIN-BINDING (POTENTIAL).  
 FT DOMAIN 379 548 VMFC.  
 FT DOMAIN 549 690 3 X TSP TYPE-1 REPEATS (CS-LIKE).  
 FT DOMAIN 723 950 7 X TSP TYPE-3 REPEATS (CA-BINDING).  
 FT DOMAIN 951 1170 C-TERMINAL.  
 FT REPEAT 351 430 TSP TYPE-1 1.  
 FT REPEAT 435 491 TSP TYPE-1 2.  
 FT REPEAT 492 548 TSP TYPE-1 3.  
 FT DOMAIN 549 587 EGF-LIKE 1.  
 FT DOMAIN 588 645 EGF-LIKE 2.  
 FT DOMAIN 646 690 EGF-LIKE 3.  
 FT REPEAT 723 758 TSP TYPE-3 1.  
 FT REPEAT 759 781 TSP TYPE-3 2.  
 FT REPEAT 782 817 TSP TYPE-3 3.  
 FT REPEAT 818 840 TSP TYPE-3 4.  
 FT REPEAT 841 878 TSP TYPE-3 5.  
 FT REPEAT 879 914 TSP TYPE-3 6.  
 FT REPEAT 915 950 TSP TYPE-3 7.  
 FT SITE 926 928 CELL ATTACHMENT SITE (POTENTIAL).  
 FT DISULFID 270 270 INTERCHAIN (PROBABLE).

```

Query Match      80.4%; Score 45; DB 1; Length 343;
Best Local Similarity 71.4%; Pred. No. 5.07e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 237 LIPTYPE 243
QY 261 LIPSPVD 267

RESULT 4
ID VL2_HPV60 STANDARD; PRT; 525 AA.
AC Q80946; AC
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE MINOR CAPSID PROTEIN L2.
OS Human papillomavirus type 60.
GN
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
RN
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U31792; AAA79490.1;
CC PFAM; PF00513; Late protein.
CC Coat protein; Late protein.
CC SEQUENCE 525 AA; 56999MW; 65477B57D988ED99 CRC64;

Query Match      80.4%; Score 45; DB 1; Length 525;
Best Local Similarity 85.7%; Pred. No. 5.07e+00;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 486 LIPNVPD 492
QY 261 LIPSPVD 267

RESULT 5
ID Y126_HUMAN STANDARD; PRT; 1073 AA.
AC Q14139; AC
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOHEILICAL PROTEIN KIAA0126.
GN KIAA0126.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN
RP SEQUENCE FROM N.A.
RC TISSUE-BONE MARROW;
RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
RX MEDLINE; 96127530.
RT "Prediction of the coding sequences of unidentified human genes. IV.
RT The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RT analysis of cDNA clones from human cell line KG-1."
RL DNA Res. 2:167-174(1995).
CC -1- SIMILARITY: TO YEAST UB FUSION DEGRADATION PROTEIN 2 (UFD2)
CC AND TO C-ELEGANS TOSH1.5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; D50916; BAA09475.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 1073 AA; 123537 MW; FBB580683DDED37D CRC64;

Query Match 80.4%; Score 45; DB 1; Length 1073;  
 Best Local Similarity 62.5%; Pred. No. 5.07e+00;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 498 LIPAVOEP 505  
 QY 261 LIPSPDP 268

RESULT 6  
 ID TSPI\_HUMAN STANDARD; PRT; 1170 AA.  
 AC P07996;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE THROMBOSPONDIN 1 PRECURSOR.  
 GN THBS1 OR TSPI OR TSPC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-ENDOTHELIAL CELLS;  
 RX MEDLINE; 87057617.  
 RA Lawler J., Hynes R.O.;  
 RT "The structure of human thrombospondin, an adhesive glycoprotein with  
 RT multiple calcium-binding sites and homologies with several different  
 RT proteins.";  
 RL J. Cell Biol. 103:1635-1648(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89139590.  
 RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,  
 RA Baumgartel D.M., Rotwein P., Frazier W.A.;  
 RT "Complete thrombospondin mRNA sequence includes potential regulatory  
 RT sites in the 3' untranslated region.";  
 RL J. Cell Biol. 108:729-736(1989).  
 RN [3]  
 RP SEQUENCE OF 1-397 FROM N.A.  
 RX MEDLINE; 87157592.  
 RA Kobayashi S., Eden-McCutchan F., Framson P., Bornstein P.;  
 RT "Partial amino acid sequence of human thrombospondin as determined by  
 RT analysis of cDNA clones: homology to malarial circumsporozoite  
 RT proteins.";  
 RL Biochemistry 25:8418-8425(1986).  
 RN [4]  
 RP SEQUENCE OF 1-374 FROM N.A.  
 RX MEDLINE; 86287276.  
 RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;  
 RT "Characterization of a cDNA encoding the heparin and collagen binding  
 RT domains of human thrombospondin.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).  
 RN [5]  
 RP SEQUENCE OF 1-166 FROM N.A.  
 RX MEDLINE; 89291870.  
 RA Laherty C.D., Gierman T.M., Dixit V.M.;  
 RT "Characterization of the promoter region of the human thrombospondin  
 RT gene. DNA sequences within the first intron increase transcription.";  
 RL J. Biol. Chem. 264:11222-11227(1989).  
 RN [6]  
 RP SEQUENCE OF 1028-1170 FROM N.A.  
 RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;  
 RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases  
 CC -!- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND  
 CC CELL-TO-MATRIX INTERACTIONS. CAN BIND TO FIBRINOGEN, FIBRONECTIN,

CC LAMININ AND TYPE V COLLAGEN.  
 CC -!- SUBUNIT: HOMOTRIMER, CROSS-LINKED BY DISULFIDE BONDS.  
 CC -!- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 VWF-C DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 3 TYPE-1 TSP REPEATS AND 7 TYPE-3 TSP REPEATS  
 CC (WHICH BIND CALCIUM).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; M25631; AAA36741.1; -  
 DR EMBL; X04665; CAA38370.1; -  
 DR EMBL; X14787; CAA32889.1; -  
 DR EMBL; J04835; AAA61178.1; -  
 DR EMBL; M99425; AAB59366.1; -  
 DR PIR; A05172; A05172.  
 DR PIR; A25812; A25812.  
 DR PIR; A26155; A26155.  
 DR PIR; A30140; A30140.  
 DR PIR; A34274; A34274.  
 DR HSSP; P35555; IEMO.  
 DR MIM; 188060;  
 DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01208; VWF\_C; 1.  
 DR PFAM; PF00008; EGF; 2.  
 DR PFAM; PF00090; tsp\_1; 3.  
 DR PFAM; PF00093; vwc; 1.  
 KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;  
 KW EGF-like domain; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 1170  
 FT DOMAIN 19 232  
 FT DOMAIN 316 373  
 FT DOMAIN 379 548  
 FT DOMAIN 549 690  
 FT DOMAIN 723 950  
 FT DOMAIN 951 1170  
 FT REPEAT 379 430  
 FT REPEAT 435 491  
 FT REPEAT 492 548  
 FT DOMAIN 549 587  
 FT DOMAIN 588 645  
 FT DOMAIN 646 690  
 FT REPEAT 723 758  
 FT REPEAT 759 781  
 FT REPEAT 782 817  
 FT REPEAT 818 840  
 FT REPEAT 841 878  
 FT REPEAT 879 914  
 FT REPEAT 915 950  
 FT SITE 926 928  
 FT DISULFID 260 270  
 FT DISULFID 274 274  
 FT DISULFID 551 562  
 FT DISULFID 556 572  
 FT DISULFID 575 586  
 FT DISULFID 592 608  
 FT DISULFID 617 644  
 FT DISULFID 620 644  
 FT DISULFID 650 663  
 FT DISULFID 657 676  
 FT DISULFID 678 689  
 FT CARBOHYD 248 248  
 FT CARBOHYD 360 360  
 FT CARBOHYD 708 708  
 FT CARBOHYD 1067 1067

THROMBOSPONDIN 1.  
 HEPARIN-BINDING (POTENTIAL).  
 VWF-C  
 3 X TSP TYPE-1 REPEATS (CS-LIKE).  
 3 X EGF-TYPE REPEATS.  
 7 X TSP TYPE-3 REPEATS (CA-BINDING).  
 C-TERMINAL.  
 TSP TYPE-1 1.  
 TSP TYPE-1 2.  
 TSP TYPE-1 3.  
 EGF-LIKE 1.  
 EGF-LIKE 2.  
 EGF-LIKE 3.  
 TSP TYPE-3 1.  
 TSP TYPE-3 2.  
 TSP TYPE-3 3.  
 TSP TYPE-3 4.  
 TSP TYPE-3 5.  
 TSP TYPE-3 6.  
 TSP TYPE-3 7.  
 CELL ATTACHMENT SITE (POTENTIAL).  
 INTERCHAIN (PROBABLE).  
 INTERCHAIN (PROBABLE).  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 POTENTIAL.  
 POTENTIAL.  
 POTENTIAL.





HEPARAN SULFATE (POTENTIAL).  
MEDIATES MOTOR NEURON ATTACHMENT  
(POTENTIAL).

Query Match 80.4%; Score 45; DB 1; Length 3707;  
Best Local Similarity 75.0%; Pred. No. 5.07e+00;

```

Matches      6; Conservative      1; Mismatches      1; Indels      0; Gaps      0;
Db      275      LLPSVPGP      282
      |:|||||
QY      261      LIPSVPOP      268

```

RESULT	8	
ID	BM3B_HUMAN	STANDARD; + PRT; 478 AA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announcement/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
DR      MIM: 601361; -.
DR      PROSITE; PS00250; TGF_BETA; 1.
DR      PFAM; PF00019; TGF-beta; 1
```

```

Query Match      78.6%      Score 44;  DB 1;  Length 478;
Best Local Similarity 50.0%;
Matches: 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

434	IIPGIP	441
:	:	:
:	:	:
:	:	:
261	LIPSPDP	268
Y		

RESULT		9		
D	BCCA MYCLE		STANDARD;	PRT; 598 AA.
C	P46392:			

US-09-376-430-2-08.isp

Thu May 11 06:49:43 2000

DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE ACETYL-PROPIONYL-COENZYME A CARBOXYLASE, ALPHA CHAIN [INCLUDES: BIOTIN  
DE CARBOXYLASE (EC 6.3.4.14); BIOTIN CARBOXYL CARRIER PROTEIN (BCPP)]  
GN BCCA OR B1308\_G1\_129.  
OC Bacteria: Firmicutes; Actinobacteria: Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE; 9422829.  
RA Norman E., de Smet K.A.L., Stoker N.G., Ratledge C., Wheeler P.R.,  
RA Dale J.W., "Lipid synthesis in mycobacteria: characterization of the biotin  
RT carboxyl carrier protein genes from Mycobacterium leprae and M.  
RT tuberculosis.", J. Bacteriol. 176:2525-2531(1994).  
RL [2]  
RN SEQUENCE FROM N.A.  
RX Smith D.R., Robinson K.  
RA Submitted (MAR1994) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: THIS PROTEIN CARRIES TWO FUNCTIONS: BIOTIN CARBOXYL  
CC CARRIER PROTEIN AND BIOTIN CARBOXYLTRANSFERASE.  
CC -1- CATALYTIC ACTIVITY: BPP + BIOTIN-CARBOXYL-CARRIER PROTEIN + CO(2)  
CC -> ADP + ORTHOPHOSPHATE + CARBOXYBIOTIN-CARBOXYL-CARRIER PROTEIN.  
CC -1- COFACTOR: BIOTIN.  
CC -1- PATHWAY: FIRST STEP IN LONG-CHAIN FATTY ACID SYNTHESIS.  
CC -1- SUBUNIT: MULTIMER COMPRSED OF 2 DIFFERENT SUBUNITS, THE LARGER  
CC ONE (63/64 KDa) HAS BIOTIN CARBOXYLASE AND BIOTIN CARRIER  
CC FUNCTION, WHILE THE SMALLER SUBUNIT POSSESSES CARBOXYLTRANSFERASE  
CC AND SUBSTRATE BINDING ACTIVITY.  
CC -1- SIMILARITY TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-  
CC PHOSPHATE SYNTHETASES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; X63470; CRA45070.1; -  
CC EMBL; U00012; CRA85920.1; -  
CC HSP; P24182; LINC; BIOTIN; 1.  
CC PROSITE; PS00188; CPSASE.1; 1.  
CC PROSITE; PS00866; CPSASE.2; 1.  
CC PROSITE; PS00867; CPSASE.2; 1.  
CC PFAM; PF00289; CPSase\_Lchain; 1.  
CC PFAM; PF00364; biotin\_lipoyl; 1.  
CC KW Fatty acid biosynthesis; Ligase; Biotin; Multifunctional enzyme;  
KW ATP-binding. 11 441 BIOTIN CARBOXYLASE.  
FT DOMAIN 532 598 BIOTIN CARBOXYL CARRIER PROTEIN.  
FT NP\_BIND 169 174 ATP (BY SIMILARITY).  
FT ACT\_SITE 299 299 BY SIMILARITY.  
FT BINDING 564 564 BIOTIN (BY SIMILARITY).  
FT CONFLICT 30 30 H -> D (IN REF. 2).  
SQ SEQUENCE 598 AA; 63885 MW; 5D1A939348EED5D CRC64;  
Query Match 78.6%; Score 44; DB 1; Length 598;  
Best Local Similarity 62.5%; Pred. No. 8.70e+00;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Db 137 LVPGTDPD 144  
QY 261 LIFSVPDP 268  
RESULT 10 STANDARD; PRT; 334 AA.  
ID EXOO\_RHIME  
AC P33697;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE SUCCINOGLUCAN BIOSYNTHESIS PROTEIN EXOO.  
GN EXOO.  
OC Rhizobium meliloti.  
OC Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
RN [1]  
RN SEQUENCE FROM N.A.  
RX STRAIN=1021;  
RX MEDLINE; 94042869.  
RA Glucksmann M.A., Reuber T.L., Walker G.C.;  
RA "Family of glycosyl transferases needed for the synthesis of  
RT succinoglycan by Rhizobium meliloti.",  
RL J. Bacteriol. 175:7033-7044(1993).  
RN [2]  
RN SEQUENCE FROM N.A.  
RX STRAIN=2011;  
RX MEDLINE; 94067019.  
RA Becker A., Kleickmann A., Keller M., Arnold W., Puehler A.;  
RA "Identification and analysis of the Rhizobium meliloti exoAMONP genes  
RT involved in exopolysaccharide biosynthesis and mapping of promoters  
RT located on the exoKLAONP fragment.",  
RL Mol. Gen. Genet. 241:367-379(1993).  
CC -1- PATHWAY: EXOPOLYSACCHARIDE BIOSYNTHESIS.  
CC -1- SIMILARITY: BELONGS TO THE AMSBE/EXOOU FAMILY OF GLYCOSYL  
CC TRANSFERASES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; L20758; AAA16044.1; -  
CC EMBL; Z22636; CAA80347.1; -  
CC PIR; C49348; C49348.  
CC PIR; S37029; S37029.  
CC PIR; S39958; S39958.  
CC PFAM; PF00535; Glycos\_transf\_2; 1.  
CC KW Exopolysaccharide synthesis; Transferase.  
CC CONFLICT 208 334  
SQ SEQUENCE 334 AA; 36481 MW; 07FEFD1BFDE68875 CRC64;  
Query Match 76.8%; Score 43; DB 1; Length 334;  
Best Local Similarity 50.0%; Pred. No. 1.48e+01;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
Db 56 LVAAIPDP 63  
QY 261 LIFSVPDP 268  
RESULT 11 STANDARD; PRT; 646 AA.  
ID IRGI\_MOUSE  
AC P54987;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE IMMUNE-RESPONSIVE PROTEIN 1 (FRAGMENT).  
GN IRGI.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C;  
RX MEDLINE; 95237894.  
RT Lee C.G.L., Jenkins N.A., Gilbert D.J., Copeland N.G., O'Brien W.E.;  
RT "Cloning and analysis of gene regulation of a novel LPS-inducible  
RT cDNA.";  
RL Immunogenetics 41:263-270(1995).  
CC -!- INDUCTION: FOLLOWING LIPOPOLYSACCHARIDE (LPS) STIMULATION.  
CC -!- SIMILARITY: TO B.SUTILIS YXEQ.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L38281; AAA74554.1; -  
DR MGD; MGI:103206; IRG1.  
FT NON\_TER 1  
SQ SEQUENCE 646 AA; 71285 MW; F55626D3B3CB4B8A CRC64;  
Query Match 76.8%; Score 43; DB 1; Length 646;  
Best Local Similarity 62.5%; Pred. No. 1.48e+01;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Db 78 LLPSTPEP 85  
QY 261 LIPSVDP 268  
:|:|:|:|  
RESULT 12  
ID INV1\_MAIZE STANDARD; PRT; 670 AA.  
AC P49175;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE BETA-FRUCTOFURANOSIDASE 1 PRECURSOR (EC 3.2.1.26) (SUCROSE-6-PHOSPHATE  
DE HYDROLASE 1) (INVERTASE 1).  
GN IVRL.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
OC Poaceae; Zea.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. B73;  
RX Xu J., Pemberton G.H., Almira E.C., McCarty D.R., Koch K.E.;  
RA "The Ivr 1 gene for invertase in maize.";  
RT Plant Physiol. 108:1293-1294(1995).  
RL -!- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING BETA-D-  
CC FRUCTOFURANOSIDE RESIDUES IN BETA-D-FRUCTOFURANOSIDES.  
CC -!- SUBCELLULAR LOCATION: VACUOLAR.  
CC -!- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U16123; AA83439.1; -  
DR MA1ZEDB; 86037; -  
DR PROSITE; PS00609; GLYCOSYL\_HYDROL\_F32; 1.  
DR PFW; PF00251; Glyco\_hydro\_32; 1.  
KW Hydrolase; Glycosidase; Glycoprotein; Zymogen; Signal;  
KW Multigene family.  
FT SIGNAL 1

FT PROPEP 2 2  
FT CHAIN 2 670  
FT ACT\_SITE 139 139 BETA-FRUCTOFURANOSIDASE 1.  
FT CARBOHYD 165 165 BY SIMILARITY.  
FT CARBOHYD 275 275 POTENTIAL.  
FT CARBOHYD 518 518 POTENTIAL.  
FT CARBOHYD 595 595 POTENTIAL.  
FT CARBOHYD 639 639 POTENTIAL.  
SQ SEQUENCE 670 AA; 71932 MW; DEDE0989C7E6AEB0 CRC64;  
Query Match 76.8%; Score 43; DB 1; Length 670;  
Best Local Similarity 62.5%; Pred. No. 1.48e+01;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
Db 1 MIPAVADP 8  
QY 261 LIPSVDP 268  
:|:|:|:|  
RESULT 13  
ID SNF5\_YEAST STANDARD; PRT; 905 AA.  
AC P18480;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE TRANSCRIPTION REGULATORY PROTEIN SNF5 (SWI/SNF COMPLEX COMPONENT SNF5)  
DE (TRANSCRIPTION FACTOR TYE4).  
GN SNF5 OR TYE4 OR SWI10 OR YBR289W OR YBR2036.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;  
OC Saccharomycetaceae; Saccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MCY;  
RX MEDLINE; 91042489.  
RA Laurent B.C., Treitel M.A., Carlson M.;  
RT "The SNF5 protein of Saccharomyces cerevisiae is a glutamine- and  
RT proline-rich transcriptional activator that affects expression of a  
RT broad spectrum of genes.";  
RL Mol. Cell. Biol. 10:5616-5625(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE; 94378722.  
RA Holmstrom K., Brandt T., Kallio T.;  
RT "The sequence of a 32,420 bp segment located on the right arm of  
RT chromosome II from Saccharomyces cerevisiae.";  
RL Yeast 10:847-862(1994).  
CC -!- FUNCTION: INVOLVED IN TRANSCRIPTIONAL ACTIVATION. THE SWI/SNF  
CC COMPLEX IS REQUIRED FOR THE INDUCED EXPRESSION OF A LARGE NUMBER  
CC OF GENES. THIS COMPLEX ALTERS CHROMATIN STRUCTURE TO FACILITATE  
CC BINDING OF GENE-SPECIFIC DEDICATED TRANSCRIPTION FACTORS.  
CC -!- SUBUNIT: COMPONENT OF THE SWI/SNF GLOBAL TRANSCRIPTION ACTIVATOR  
CC COMPLEX.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M36482; AAA85062.1; -  
DR EMBL; X76053; CAA53652.1; -  
DR EMBL; Z36158; CAA85254.1; -  
DR PIR; S44531; RGYBS.  
DR PIR; S39145; S39145.  
DR SGD; L0001948; SNF5.  
KW Transcription regulation; Activator; Nuclear protein.  
FT DOMAIN 31 270 GLN-RICH.  
FT DOMAIN 72 132 PRO-RICH.

```

FT DOMAIN 272 324 PRO-RICH.
FT DOMAIN 489 588 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 714 882 PRO-RICH.
FT DOMAIN 755 798 ARG/LYS-RICH (BASIC).
FT CONFLICT 564 564 E->D (IN REF. 1).
SQ SEQUENCE 905 AA; 102557 MW; A287B4A648DDIA35 CRC64;
Query Match 76.88; Score 43; DB 1; Length 905;
Best Local Similarity 71.4; Pred. No. 1.48e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 836 IPSIPNP 842
QY 262 IPSVPDP 268

RESULT 14
ID DNJM_MYCPN STANDARD; PRT; 910 AA.
AC P75354;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DNJ-LIKE PROTEIN MG200 HOMOLOG.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE; 97105885.
RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Hermann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SIMILARITY: CONTAINS A DNJ-LIKE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; A5000004; AAB95683.1; -
CC HSSP; P25685; 1HDJ.
CC PROSITE; PS00636; DNJ_1; 1.
CC PROSITE; PS50076; DNJ_2; 1.
CC PFAM; PF00226; Dnaj; 1.
CC KW Hypothetical protein; Chapertone.
FT DOMAIN 4 73 DNJ-LIKE.
SQ SEQUENCE 910 AA; 100190 MW; 125D0E37D2D221A7 CRC64;
Query Match 76.88; Score 43; DB 1; Length 910;
Best Local Similarity 57.18; Pred. No. 1.48e+01;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 634 LVTPE 640
QY 261 LIPSVPD 267

RESULT 15
ID EVGS_ECOLI STANDARD; PRT; 1197 AA.
AC P30855; P77644;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PUTATIVE SENSOR PROTEIN EVGS PRECURSOR (EC 2.7.3.-).
GN EVGS.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

```

```

OC Escherichia.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE; 94171083.
RA Utsumi R., Katayama S., Taniguchi M., Horie T., Ikeda M., Igaki S.,
RA Nakagawa H., Miwa A., Tanabe H., Noda M.;
RT "Newly identified genes involved in the signal transduction of
RL Escherichia coli K-12.";
Gene 140:73-77(1994).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE; 93173621.
RA Utsumi R., Katayama S., Ikeda M., Igaki S., Nakagawa H., Miwa A.,
RA Taniguchi M., Noda M.;
RT "Cloning and sequence analysis of the evgAS genes involved in signal
RL transduction of Escherichia coli K-12.";
Nucleic Acids Symp. Ser. 422:149-150(1992).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE; 97426617.
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
Science 277:1453-1474(1997).
[4]
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Alba H., Baba T., Fujita K., Hayashi K., Honjo A., Horiuchi T.,
RA Ikemoto K., Inada T., Isono S., Itoh T., Kanai K., Kasai H.,
RA Kashimoto K., Kim S., Kimura S., Kitagawa M., Kitakawa M., Makino K.,
RA Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y.,
RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM EVGS/EVGA.
CC EVGS MAY FUNCTION AS A MEMBRANE-ASSOCIATED PROTEIN KINASE THAT
CC PHOSPHORYLATES EVGA IN RESPONSE TO ENVIRONMENTAL SIGNALS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (PROBABLE).
CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D14008; BAA03108.1; -
CC EMBL; A6000325; AAC75429.1; -
CC EMBL; D90867; CAB22163.1; -
CC PIR; J00221; J00221.
CC HSSP; P06143; 1UDR.
CC ECOGENE; EG11610; EVGS.
CC PFAM; PF00072; response_reg; 1.
CC PFAM; PF00512; signal; 1.
CC PFAM; PF01627; Hpt; 1.
CC
CC Sensory transduction; Transferase; Kinase; Phosphorylation;
KW Transmembrane; Inner membrane; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 1197 PUTATIVE SENSOR PROTEIN EVGS.
FT DOMAIN 22 378 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 379 394 POTENTIAL.
FT DOMAIN 395 535 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 536 552 POTENTIAL.
FT DOMAIN 553 1197 CYTOPLASMIC (POTENTIAL).

```

FT MOD\_RES 721 721 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT CONFLICT 152 152 L -> F (IN REF. 1 AND 2).  
FT CONFLICT 242 243 EF -> PL (IN REF. 1 AND 2).  
FT CONFLICT 275 275 W -> R (IN REF. 1 AND 2).  
FT CONFLICT 420 421 SQ -> FE (IN REF. 1 AND 2).  
FT CONFLICT 739 739 G -> D (IN REF. 1 AND 2).  
FT CONFLICT 758 758 G -> K (IN REF. 1 AND 2).  
FT CONFLICT 761 761 L -> V (IN REF. 1 AND 2).  
FT CONFLICT 877 877 S -> L (IN REF. 1 AND 2).  
FT CONFLICT 1045 1045 R -> H (IN REF. 1 AND 2).  
FT CONFLICT 1074 1074 H -> Y (IN REF. 1 AND 2).  
SQ SEQUENCE 1197 AA; 134742 MW; E8E1DE0F797B1278 CRC64;  
Query Match 76.8%; Score 43; DB 1; Length 1197;  
Best Local Similarity 62.5%; Pred. No. 1.48e+01;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
Db 475 LIPGVPNA 482  
Qy 261 LIPSVDP 268  
|||||:

Search completed: Wed May 10 12:04:33 2000  
Job time : 102 secs.

\*\*\*\*\*  
 W P S R L H  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Blocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 12:04:51 2000; MasPar time 224.17 Seconds  
 Tabular output not generated. 2.474 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (261-268) from US9376430A.pep (8 of 25)  
 Perfect Score: 56  
 Sequence: 1 LIPSVDPDP 8

Scoring table: PAM 150  
 Gap 11

Searched: 225878 seqs, 69334122 residues  
 Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: spiremb12  
 1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
 5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
 9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
 13:sp-vertebrate 14:sp-virus

Statistics: Mean 21.255; Variance 24.636; scale 0.863

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	47	83.9	782	4	PDE4C-791 (FRAGMENT)	7.59e+00
2	47	83.9	791	4	PDE4C-791 (PDE4C-426)	7.59e+00
3	46	82.1	244	2	HYPOTHETICAL 27.2 KD P	1.26e+01
4	46	82.1	258	2	IS600 PROTEIN (FRAGMEN	1.26e+01
5	46	82.1	325	2	OLIGOPEPTIDE ABC TRANS	1.26e+01
6	46	82.1	728	2	ABC EXCISION NUCLEASE	1.26e+01
7	45	80.4	229	6	THROMBOSPONDIN-1 (FRAG	2.06e+01
8	45	80.4	243	6	THROMBOSPONDIN 23	2.06e+01
9	45	80.4	297	10	MITSUGUMIN 23	2.06e+01
10	45	80.4	320	1	T32G6.18 PROTEIN	2.06e+01
11	45	80.4	320	3	320AA LONG HYPOTHETICA	2.06e+01
12	45	80.4	356	2	MAT-2 GENE	2.06e+01
13	45	80.4	376	4	D.D CARBOXYPEPTIDASE	2.06e+01
14	45	80.4	378	2	THROMBOSPONDIN (TSP) P	2.06e+01
15	45	80.4	419	2	PUTATIVE GLYCOSYLTRANS	2.06e+01
16	45	80.4	479	5	HYPOTHETICAL 38.8 KD P	2.06e+01
17	45	80.4	562	3	DP2W10 (FRAGMENT)	2.06e+01
18	45	80.4	623	10	DP2W10 (FRAGMENT)	2.06e+01
19	45	80.4	1120	4	MATING TYPE PROTEIN MA	2.06e+01
20	45	80.4	1170	6	ENDO-BETA-1,4-D-GLUCAN	2.06e+01
					TRANSCRIPTIONAL INTERM	2.06e+01
					THROMBOSPONDIN 1	2.06e+01

21 45 80.4 1708 14 Q9YLRI  
 22 44 78.6 142 1 Q9YDD3  
 23 44 78.6 278 2 P71053  
 24 44 78.6 391 10 Q22228  
 25 44 78.6 399 4 Q9Y5P5  
 26 44 78.6 539 10 Q04422  
 27 44 78.6 600 2 Q96890  
 28 44 78.6 609 2 Q923X7  
 29 44 78.6 1767 5 Q24495  
 30 43 76.8 39 5 Q46984  
 31 43 76.8 229 2 Q94406  
 32 43 76.8 345 14 Q41524  
 33 43 76.8 386 1 Q9YFS3  
 34 43 76.8 388 13 Q9W638  
 35 43 76.8 389 2 Q9Z164  
 36 43 76.8 443 3 Q02398  
 37 43 76.8 477 3 Q00183  
 38 43 76.8 628 10 Q04831  
 39 43 76.8 1077 10 Q81441  
 40 43 76.8 5825 10 Q82731  
 41 42 75.0 187 1 Q9YD91  
 42 42 75.0 333 1 Q9YC52  
 43 42 75.0 455 5 Q96838  
 44 42 75.0 1402 11 Q920W3  
 45 42 75.0 2150 5 Q17596

ALIGNMENTS

RESULT 1  
 ID 076105 PRELIMINARY; PRT; 782 AA.  
 AC 076105:  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
 DE PDE4C-791 (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LAMERDIN J.E., MCCREADY P.M., SKOWRONSKI E., VISWANATHAN V.,  
 RA BURKHART-SCHULTZ K., GORDON L., DIAS J., RAMIREZ M., STILWAGEN S.,  
 RA PHAN H., VELASCO N., DO L., REGALA W., TERRY A., GARNES J.,  
 RA DANCANAN L., FOUNDSTONE P., CHRISTENSEN M., GEORGESCU A., AVILA J.,  
 RA LIU S., ATTIX C., ANDREISE T., FRANKHEIM M., AMICO-KELLER G.,  
 RA COFIELD J., DUARTE S., LUCAS S., BRUCE R., THOMAS P., QUAN G.,  
 RA KRONMILLER B., ARELLANO A., SAUNDERS C., OW D., NOLAN M., TRONG S.,  
 RA KOBAYASHI A., OLSEN A.S., CARRANO A.V.,  
 RT "sequence analysis of an -600 kb region in 19p13.1 between JAK3 and  
 RT PDE4C".  
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AC005759; AAC83050.1;  
 DR PROSITE; PS00126; PDEASE\_1;  
 FT NON-TER.  
 FT SEQUENCE 782 AA; 87342 MW; CB3754C4 CRC32;  
 SQ  
 Query Match 83.9%; Score 47; DB 4; Length 782;  
 Best Local Similarity 62.5%; Pred. No. 7.59e+00;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 DB 59 LIPSVDPDP 66  
 QY 261 LIPSVDPDP 268.  
 RESULT 2  
 ID 043849 PRELIMINARY; PRT; 791 AA.  
 AC 043849:  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)

DE PDE4C-791 (PDE4C-426) (FRAGMENT).  
 GN PDE4C.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 RN Eutheria; Primates; Catarrhini; Homidae; Homo.  
 RC [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE-TESTIS;  
 RX MEDLINE; 95145731.  
 RA ENGELS P., SULLIVAN M., MULLER T., LUBBERT H.;  
 RT "Molecular cloning and functional expression in yeast of a human CAMP-  
 RT specific phosphodiesterase subtype (PDE IV-C).";  
 RL FEBS Lett. 358:305-310(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE-TESTIS, AND LUNG;  
 RX MEDLINE; 98007880.  
 RA OBERNOLTE R., RATLIFF J., BAECKER P.A., DANIELS D.V., ZUPPAN P.,  
 RA JARNAGIN K., SHELTON E.R.;  
 RT "Multiple splice variants of phosphodiesterase PDE4C cloned from human  
 RT lung and testis";  
 RL Biochim. Biophys. Acta 1353:287-297(1997).  
 DR EMBL; U66346; AAB96875.1; -.  
 DR EMBL; U66347; AAB96876.1; -.  
 DR PROSITE; PS00126; PDEASE\_I; 1.  
 DR PFAM; PF00233; PDEase; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 791 AA; 88142 MW; AFDA976A CRC32;

Query Match 83.9%; Score 47; DB 4; Length 791;  
 Best Local Similarity 62.5%; Pred. No. 7.59e+00;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Db 68 LLPAPDP 75  
 QY 261 LIPSVDP 268  
 :|:|:|

RESULT 3  
 ID Q9ZGV2 PRELIMINARY; PRT; 244 AA.  
 AC Q9ZGV2;  
 DT 01-MAY-1999 (TRENBLrel. 10, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DE 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
 GN HYPOTHETICAL 27.2 KD PROTEIN.  
 OS Escherichia coli O157:H7.  
 OG Plasmid poi57.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-EDL933;  
 RX MEDLINE; 98391744.  
 RA BURLAND V., SHAO Y., PERNA N.T., PLUNKETT G., SOFIA H.J.,  
 RA BLATTNER F.R.;  
 RT "The complete DNA sequence and analysis of the large virulence plasmid  
 RT of Escherichia coli O157:H7";  
 RL Nucleic Acids Res. 26:4196-4204(1998).  
 DR EMBL; AF074613; AAC70082.1; -.  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 244 AA; 27242 MW; 06F395F1 CRC32;

Query Match 82.1%; Score 46; DB 2; Length 244;  
 Best Local Similarity 62.5%; Pred. No. 1.26e+01;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Db 26 LVPGVDP 33  
 QY 261 LIPSVDP 268  
 :|:|:|

RESULT 4

ID O82931 PRELIMINARY; PRT; 258 AA.  
 AC O82931;  
 DT 01-NOV-1998 (TRENBLrel. 08, Created)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
 DE IS600 PROTEIN (FRAGMENT).  
 GN IS600.  
 OS Escherichia coli.  
 OG Plasmid poi57.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 RN Escherichia.  
 RC [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-O157:H7;  
 RX MEDLINE; 98290540.  
 RA MAKINO K., ISHII K., YASUNAGA T., HATTORI M., YOKOYAMA K.,  
 RA YATSUDO H.C., KUBOTA Y., YAMAICHI Y., IIDA T., YAMAMOTO K., HONDA T.,  
 RA HAN C., OHTSUBO A., KASAMATSU M., HAYASHI T., KUHARA S., SHINAGAWA H.;  
 RT "Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an  
 RT enterohemorrhagic Escherichia coli O157:H7 derived from Sakai  
 RT outbreak";  
 RL DNA Res. 5:1-9(1998).  
 DR EMBL; AB011549; BAA31830.1; -.  
 KW Plasmid.  
 FT NON\_TER 1  
 SQ SEQUENCE 258 AA; 28805 MW; 13405C00 CRC32;

Query Match 82.1%; Score 46; DB 2; Length 258;  
 Best Local Similarity 62.5%; Pred. No. 1.26e+01;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Db 40 LVPGVDP 47  
 QY 261 LIPSVDP 268  
 :|:|:|

RESULT 5  
 ID Q9XON5 PRELIMINARY; PRT; 325 AA.  
 AC Q9XON5;  
 DT 01-NOV-1999 (TRENBLrel. 12, Created)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)  
 DE OLIGOPETIDE ABC TRANSPORTER, ATP-BINDING PROTEIN.  
 GN TM1151.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogales; Thermotoga.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 99287316.  
 RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,  
 RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,  
 RA McDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,  
 RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,  
 RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,  
 RA SMITH H.O., VENTER J.C., FRASER C.M.;  
 RT "Evidence for lateral gene transfer between Archaea and bacteria from  
 RT genome sequence of Thermotoga maritima";  
 RL Nature 399:323-329(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,  
 RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,  
 RA McDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,  
 RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,  
 RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,  
 RA SMITH H.O., VENTER J.C., FRASER C.M.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE001772; AAD36227.1;  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 325 AA; 37193 MW; 2BFC6A78 CRC32;

Query Match 82.1%; Score 46; DB 2; Length 325;



Best Local Similarity 62.5%; Pred. No. 1.26e+01;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 259 LISAVPDP 266  
||:|||||  
QY 261 LIPSVDP 268

RESULT 6  
ID Q9Z512 PRELIMINARY; PRT; 728 AA.

AC Q9Z512;  
DT 01-MAY-1999 (TEMBLrel. 10, Created)  
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TEMBLrel. 10, Last annotation update)  
DE ABC EXCISION NUCLEASE SUBUNIT C.  
GN SC54.13C.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA SEEGER K., HARRIS D.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA BENTLEY S.D., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RX MEDLINE; 97000351.  
RA REDENBACH M., KIESER H.M., DENAPAITE D., EICHNER A., CULLUM J.,  
RA KINASHI H., HOPWOOD D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
the 8 Mb streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
DR EMBL; AL035591; CAB38143.1; -.  
SQ SEQUENCE 728 AA; 80841 MW; A9285AF0 CRC32;

Query Match 82.1%; Score 46; DB 2; Length 728;  
Best Local Similarity 62.5%; Pred. No. 1.26e+01;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 322 LVPALPDP 329  
||:|||||  
QY 261 LIPSVDP 268

RESULT 7  
ID Q28194 PRELIMINARY; PRT; 229 AA.

AC Q28194;  
DT 01-NOV-1996 (TEMBLrel. 01, Created)  
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
DE THROMBOSPONDIN-1 (FRAGMENT).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;  
OC Bovinae; Bos.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96331130.  
RA LAFUILLADE B., PELLERIN S., KERAMIDAS M., DANIK M., CHAMBAZ E.M.,  
RA FEIGE J.J.;  
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced  
secreted protein/thrombospondin-2 expression by adrenocorticotrophic  
hormone in adrenocortical cells."  
RL J. Cell. Physiol. 167:164-172(1996).  
DR EMBL; X89511; CAA61682.1; -.  
FT NON\_TER 1  
ET NON\_TER 229 229

SQ SEQUENCE 229 AA; 25015 MW; 1DAC355E CRC32;

Query Match 80.4%; Score 45; DB 6; Length 229;  
Best Local Similarity 85.7%; Pred. No. 2.06e+01;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 48 LIPVPDP 54  
||:|||||  
QY 261 LIPSVDP 267

RESULT 8

ID 077751 PRELIMINARY; PRT; 243 AA.  
AC 077751;  
DT 01-NOV-1998 (TEMBLrel. 08, Created)  
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)  
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)  
DE MITSUGUMIN 23.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-SKELETAL MUSCLE;  
RX MEDLINE; 98385877.  
RA NISHI M., KOMAZAKI S., IINO M., KANAGAWA K., TAKESHIMA H.;  
RT "Mitsugumin23, a novel transmembrane protein on endoplasmic reticulum  
and nuclear membranes."  
RL FEBS Lett. 432:191-196(1998).  
DR EMBL; AB013721; BAA33366.1; -.  
SQ SEQUENCE 243 AA; 26145 MW; 8A9F9AD8 CRC32;

Query Match 80.4%; Score 45; DB 6; Length 243;  
Best Local Similarity 75.0%; Pred. No. 2.06e+01;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 175 LRVSVDP 182  
||:|||||  
QY 261 LIPSVDP 268

RESULT 9

ID 022227 PRELIMINARY; PRT; 297 AA.  
AC 022227;  
DT 01-JAN-1998 (TEMBLrel. 05, Created)  
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
DT 01-AUG-1998 (TEMBLrel. 07, Last annotation update)  
DE T32G6.18 PROTEIN.  
GN T32G6.18.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
OC Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLOMBIA;  
RA ROUNSLEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,  
RA SKES S.M., KAUL S., MASON T.M., KERLAVAGE A.R., ADAMS M.D.,  
RA SOMERVILLE C.R., VENTER J.C.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC002510; AAB84348.1; -.  
SQ SEQUENCE 297 AA; 32510 MW; E2AE4E29 CRC32;

Query Match 80.4%; Score 45; DB 10; Length 297;  
Best Local Similarity 71.4%; Pred. No. 2.06e+01;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 40 LIPSIPE 46  
||:|||||  
QY 261 LIPSVDP 267

RESULT 10  
 ID Q9Y9W7 PRELIMINARY; PRT; 320 AA.  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE 320AA LONG HYPOTHETICAL PEPTIDE TRANSPORT ATP-BINDING PROTEIN.  
 GN APE2261.  
 OS Aeropyrum pernix.  
 OC Archaea; Crenarchaeota; Aeropyrum.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K1;  
 RX MEDLINE; 99310339.  
 RA KAWARABAYASHI Y., HINO Y., HORIKAWA H., YAMAZAKI S., HAIKAWA Y.,  
 RA JIN-NO K., TAKAHASHI M., SEKINE M., BABA S., ANKAI A., KOSUGI H.,  
 RA HOSOIYAMA A., FUKUI S., NAGAI Y., NISHIJIMA K., NAKAZAWA H.,  
 RA TAKAMIYA M., MASUDA S., FUNAHASHI T., TANAKA T., KUDOH Y.,  
 RA YAMAZAKI J., KUSHIDA N., OGUCHI A., AOKI K., KUBOTA K., NAKAMURA Y.,  
 RA NOMURA N., SAKO Y., KIKUCHI H.;  
 RT "Complete genome sequence of an aerobic hyper-thermophilic  
 RT crenarchaeon, Aeropyrum pernix K1.";  
 RL DNA Res. 6:83-101(1999).  
 DR EMBL; AP000064; BAA81273.1; -.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 320 AA; 34740 MW; 2930B7A5 CRC32;

Query Match 80.4%; Score 45; DB 1; Length 320;  
 Best Local Similarity 87.5%; Pred. No. 2.06e+01;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 256 LIKSVDP 263  
 QY 261 LIPSVDP 268  
 |||||||

RESULT 11  
 ID Q00307 PRELIMINARY; PRT; 343 AA.  
 AC Q00307;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE MAT-2 GENE.  
 GN MAT-2.  
 OS Bipolaris sacchari.  
 OC Eukaryota; Fungi; Ascomycota; Euascomycetes; Loculoascomycetes;  
 OC Pleosporales; Pleosporaceae; anamorphic Pleosporaceae; Bipolaris.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=764-1;  
 RX MEDLINE; 96213991.  
 RA SHARON A., YAMAGUCHI K., CHRISTIANSEN S.K., HORWITZ B.A., YODER O.C.,  
 RA TURGEON B.G.;  
 RT "An asexual fungus has the potential for sexual development.";  
 RL Mol. Gen. Genet. 251:60-68(1996).  
 DR EMBL; X95814; CAA65081.1; -.  
 DR PFAM; PF00505; HMG\_box; 1.  
 KW DNA-binding.  
 SQ SEQUENCE 343 AA; 39257 MW; 477BB76D CRC32;

Query Match 80.4%; Score 45; DB 3; Length 343;  
 Best Local Similarity 71.4%; Pred. No. 2.06e+01;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 237 LIPTVE 243  
 QY 261 LIPSVDP 267  
 |||:|:

RESULT 12  
 ID Q9XCR6 PRELIMINARY; PRT; 356 AA.  
 AC Q9XCR6;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE D.D. CARBOXYPEPTIDASE.  
 GN VANDY.  
 OS Enterococcus faecium (Streptococcus faecium).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;  
 OC Enterococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BM4339;  
 RX MEDLINE; 99296571.  
 RA CASADEWALL B., COURVALIN P.;  
 RT "Characterization of the vand glycopeptide resistance gene cluster  
 RT from Enterococcus faecium BM4339.";  
 RL J. Bacteriol. 181:3644-3648(1999).  
 DR EMBL; AF130997; AAD42182.1; -.  
 KW Carboxypeptidase.  
 SQ SEQUENCE 356 AA; 38710 MW; 704F28B9 CRC32;

Query Match 80.4%; Score 45; DB 2; Length 356;  
 Best Local Similarity 71.4%; Pred. No. 2.06e+01;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 80 VPSVPEP 86  
 QY 262 IPSVDP 268  
 :|||:

RESULT 13  
 ID Q15667 PRELIMINARY; PRT; 376 AA.  
 AC Q15667;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE THROMBOSPONDIN (TSP) PRECURSOR (FRAGMENT).  
 GN THBS1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 86287276.  
 RA DIXIT V.M., HENNESSY S.W., GRANT G.A., ROTWEIN P., FRAZIER W.A.;  
 RT "Characterization of a cDNA encoding the heparin and collagen binding  
 RT domains of human thrombospondin.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92278426.  
 RA BURLIN T.R., BARNES T.M.;  
 RT "Introns in sequence tags.";  
 RL Nature 357:367-368(1992).  
 DR EMBL; M14326; AAA61237.1; -.  
 DR PROSITE; PS01208; VWFC; 1.  
 DR PFAM; PF00093; vwc; 1.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 >376 POTENTIAL.  
 FT NON\_TER 376 376  
 SQ SEQUENCE 376 AA; 41329 MW; 7956661F CRC32;

Query Match 80.4%; Score 45; DB 4; Length 376;  
 Best Local Similarity 85.7%; Pred. No. 2.06e+01;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 66 LIPVDP 72  
 QY 261 LIPSVDP 267  
 |||:|:

RESULT 14  
 ID P95720 PRELIMINARY; PRT; 379 AA.  
 AC P95720;

DT 01-MAY-1997 (TReMBLrel. 03, Created)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
 DT 01-MAY-1997 (TReMBLrel. 03, Last annotation update)  
 DE PUTATIVE GLYCOSYLTRANSFERASE (FRAGMENT).  
 OS Streptomyces capreolus  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NRRL3817;  
 RA BARRASA M.I., TERCERO J.A., JIMENEZ A.;  
 RL Eur. J. Biochem. 0:0-0(0).  
 DR EMBL; Y11036; CAA71929.1; -.  
 KW Transferase.  
 FT NON\_TER 379 379  
 SQ SEQUENCE 379 AA; 40471 MW; 09E51052 CRC32;

Query Match 80.4%; Score 45; DB 2; Length 379;  
 Best Local Similarity 71.4%; Pred.No. 2.06e+01;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 195 LIPSIPSE 201  
 QY 261 LIPSPVD 267

RESULT 15  
 ID C05589 PRELIMINARY; PRT; 419 AA.  
 AC C05589;  
 DT 01-JUL-1997 (TReMBLrel. 04, Created)  
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE HYPOTHETICAL 38.8 KD PROTEIN.  
 GN RV1004C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RX MEDLINE; 98295987.  
 RA COLE S.T., BROSCHE R., PARKHILL J., GARNIER T., CHURCHER C., HARRIS D.,  
 RA GORDON S.V., EIGLMEIER K., GAS S., BARRY III C.E., TEKAIA F.,  
 RA BADCOCK K., BASHAM D., BROWN D., CHILLINGWORTH T., CONNOR R.,  
 RA DAVIES R., DEVLIN K., FELTHAM T., GENTLES S., HAMLIN N., HOLROYD S.,  
 RA HORNSBY T., JAGELS K., KROGH A., MCLEAN J., MOULE S., MURPHY L.,  
 RA OLIVER S., OSBORNE J., QUAIL M.A., RAJANDREAM M.A., ROGERS J.,  
 RA RUTHER S., SEEGER K., SRELTON S., SQUARES S., SQUARES R., SULSTON J.E.,  
 RA TAYLOR K., WHITEHEAD S., BARRELL B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RA PARKHILL J.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z94752; CAB08158.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 419 AA; 38785 MW; 4B1ADB2D CRC32;

Query Match 80.4%; Score 45; DB 2; Length 419;  
 Best Local Similarity 85.7%; Pred.No. 2.06e+01;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 161 IPGVDP 167  
 QY 262 IPSPVD 268

Search completed: Wed May 10 12:08:46 2000  
 Job time : 235 secs.



\*\*\*\*\*  
 W P E R E H (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 12:17:41 2000; MasPar time 2.78 Seconds  
 Tabular output not generated. 68.083 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (22-29) from US09376430A.pap (9 of 25)  
 Perfect Score: 49  
 Sequence: 1 GGGGAEG 8

Scoring table: PAM 150  
 Gap 11

Searched: 189963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: a-geneseq35  
 1:geneseqp

Statistics: Mean 13.160; Variance 40.393; scale 0.326

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	ID	Description	Pred. No.
1	42	85.7	54	1	1.71e+02
2	41	83.7	412	1	2.20e+02
3	41	83.7	818	1	2.20e+02
4	41	83.7	1896	1	2.20e+02
5	41	83.7	3119	1	2.20e+02
6	40	81.6	220	1	2.20e+02
7	39	79.6	381	1	2.83e+02
8	39	79.6	436	1	3.64e+02
9	39	79.6	464	1	3.64e+02
10	39	79.6	472	1	3.64e+02
11	39	79.6	563	1	3.64e+02
12	39	79.6	800	1	3.64e+02
13	39	79.6	1717	1	3.64e+02
14	39	79.6	1831	1	3.64e+02
15	38	77.6	93	1	4.66e+02
16	38	77.6	101	1	4.66e+02
17	38	77.6	147	1	4.66e+02
18	38	77.6	272	1	4.66e+02
19	38	77.6	287	1	4.66e+02
20	38	77.6	287	1	4.66e+02
21	38	77.6	373	1	4.66e+02
22	38	77.6	413	1	4.66e+02
23	38	77.6	545	1	4.66e+02

24	38	77.6	618	1	Y06980	Recombinant pGO-11CKS	4.66e+02
25	38	77.6	873	1	Y06983	Amino acid sequence of	4.66e+02
26	38	77.6	1060	1	W23975	Homo sapiens 20q13 amp	4.66e+02
27	38	77.6	2189	1	R05222	Antigen GX5401FL encod	4.66e+02
28	38	77.6	4987	1	R10834	Rianodin receptor.	4.66e+02
29	37	75.5	11	1	W00852	MAP2 phosphopeptide II	5.96e+02
30	37	75.5	77	1	W67493	gd-tag for fusion to g	5.96e+02
31	37	75.5	388	1	Y04999	Mycobacterium species	5.96e+02
32	37	75.5	453	1	R47872	Enzyme/biocatalyst whi	5.96e+02
33	37	75.5	453	1	W97237	An enzyme capable of d	5.96e+02
34	37	75.5	453	1	W89241	Rhodococcus sp. strain	5.96e+02
35	37	75.5	453	1	W09393	Rhodococcus rhodochrou	5.96e+02
36	37	75.5	584	1	W72152	HSV-2 strain SB5 Conti	5.96e+02
37	37	75.5	5035	1	R25450	MH mutant porcine ryan	5.96e+02
38	37	75.5	5072	1	R11510	Ryanodine receptor ded	5.96e+02
39	36	73.5	160	1	Y04851	Mycobacterium species	7.61e+02
40	36	73.5	164	1	Y04852	Mycobacterium species	7.61e+02
41	36	73.5	344	1	W06600	Hyper-sensitive respons	7.61e+02
42	36	73.5	409	1	W75700	Vpr protein binding HH	7.61e+02
43	36	73.5	506	1	W75451	Human wild type tub pr	7.61e+02
44	36	73.5	924	1	W60161	Human intercellular ad	7.61e+02
45	36	73.5	1445	1	W93395	O. longistaminata Xa21	7.61e+02

## ALIGNMENTS

RESULT 1  
 ID R14371 standard; Peptide: 54 AA.  
 AC R14371;  
 DT 30-JAN-1992 (first entry)  
 DE Crystalline domain of spider silk protein.  
 KW Textiles; crystalline; tensile strength; SSP.  
 OS Nepheila clavipes.  
 PN W09116351-A.  
 PD 31-OCT-1991.  
 PF 29-MAR-1991; U02222.  
 PR 19-APR-1990; US-511114.  
 PA (USSA) US SEC OF THE ARMY.  
 PI Lombardi SJ Kaplan DL;  
 DR WPI 91-339781/46.  
 DR N-FSDB: Q15395.  
 PT New recombinant spider silk protein(s), for textiles - with  
 PT modified relative and total amts. of crystalline and amorphous  
 PT domains, for improved commercial properties, e.g tensile strength.  
 PS Claim 5; Fig.4; 46pp; English.  
 CC The sequence is a crystalline domain of spider dragline silk protein  
 CC from the golden orb-weaving spider, N. clavipes. The dragline silk  
 CC is produced in the major ampullate gland and is the strongest of the  
 CC 5-7 different silk proteins produced by the spider. The gene encod-  
 CC ing the protein can be used to express recombinant silk protein for  
 CC use in textiles, and composite and ballistic materials. Modifica-  
 CC tions of the relative amts. of amorphous and crystalline domains can  
 CC be used to alter the properties of the material.  
 CC See also R14368-R14370.  
 SQ Sequence 54 AA;

Query Match 85.7%; Score 42; DB 1; Length 54;  
 Best Local Similarity 75.0%; Pred. No. 1.71e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Db 36 GEGGAGEG 43  
 QY 22 GGGGAEG 29  
 RESULT 2  
 ID W94998 standard; Protein; 412 AA.  
 AC W94998;  
 DT 12-MAY-1999 (first entry)  
 DE Serine/threonine protein kinase-HTLAR33 (EST derived sequence).  
 KW YAK-1; serine-threonine protein kinase; HTLAR33; bone loss; ARDS;  
 KW inflammatory disorder; osteoporosis; Adult Respiratory Disease Syndrome;  
 KW arthritis; psoriasis; dermatitis; infection; HIV-1;

KW HIV-2; cachexia; immunodeficient disorder; septic shock; pain; injury;  
 KW cancer; anorexia; bulimia; Parkinson's disease; cardiovascular disease;  
 KW neurological disorder; Huntington's disease; gene therapy; gene mapping;  
 KW Gilles de la Tourette's syndrome.  
 OS Homo sapiens.  
 PN EP-894863-A1.  
 PD 03-FEB-1999.  
 PF 21-JUL-1998; 305794.  
 PR 20-FEB-1998; US-027064.  
 PR 28-JUL-1997; US-053924.  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PI Bergsma DJ, Shabon U;  
 DR WPI: 99-108353/10.  
 DR N-PSDB: X1775.  
 PT New serine-threonine kinase (HTLAR33) polypeptides and  
 PT polynucleotides - useful as diagnostic reagents and for prevention  
 PT and treatment of bone loss, neurological and inflammatory disorders  
 PT and cancer, HIV infections and angina pectoris  
 PS Claim 14; Page 17; 30pp; English.  
 CC The invention relates to a YAK-1 related serine-threonine protein kinase  
 CC polypeptide (HTLAR33). Host cells transformed with an expression vector  
 CC comprising the HTLAR33 nucleic acid are used for the recombinant  
 CC production of the protein. HTLAR33 polynucleotides and polypeptides are  
 CC useful for are useful for diagnosing susceptibility to diseases and for  
 CC screening for antagonists, agonists. These can be used in treatment to  
 CC enhance (agonist) or block (antagonist) HTLAR33 activity. Diseases  
 CC diagnosed, prevented or treated include: bone loss and inflammatory  
 CC disorders including osteoporosis, Adult Respiratory Disease Syndrome  
 CC (ARDS), Rheumatoid arthritis, osteoarthritis, inflammatory bowel disease  
 CC (IBD), psoriasis, dermatitis, asthma, allergies; bacterial, fungal,  
 CC protozoan and viral infections, especially those caused by HIV-1 or  
 CC HIV-2; HIV-associated cachexia and other immunodeficient disorders;  
 CC septic shock; pain; injury; cancers; anorexia; bulimia; Parkinson's  
 CC disease; cardiovascular disease including restenosis, atherosclerosis,  
 CC acute heart failure, myocardial infarction; hypotension; hypertension;  
 CC urinary retention; angina pectoris; ulcers; benign prostatic hypertrophy;  
 CC and psychotic and neurological disorders, including schizophrenia, manic  
 CC depression, anxiety; delirium, dementia, severe mental retardation and  
 CC dyskinesias, such as Huntington's disease or Gilles de la Tourette's  
 CC syndrome. HTLAR33 polypeptides are also useful for mapping genes to  
 CC chromosomes, allowing gene inheritance to be studied through linkage  
 CC analysis.  
 SQ Sequence 412 AA;  
 Query Match 83.7%; Score 41; DB 1; Length 412;  
 Best Local Similarity 75.0%; Pred. No. 2.20e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Db 357 GGGGGGEG 364  
 QY 22 GGGGAAG 29  
 RESULT 3  
 ID W72035 standard; Protein; 818 AA.  
 AC W72035;  
 DT 07-DEC-1998 (first entry)  
 DE HSV-2 strain SB5 Contig ID 104 ORF#3 protein.  
 KW HSV-2 strain SB5; immunological response induction; therapy;  
 KW antiviral identification; viral protein inhibitor.  
 OS Herpes simplex virus type 2.  
 PN W09820016-A1.  
 PD 14-MAY-1998.  
 PF 31-OCT-1997; U20016.  
 PR 09-JUN-1997; US-049018.  
 PR 04-NOV-1996; US-030279.  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PI Chan JY, Dabrowski-Anaral CE, Delvecchio AM, Dillon SB,  
 PI Esser KM, Leary JJ;  
 DR WPI: 98-286847/25.  
 DR N-PSDB: V62134.  
 PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention  
 PT and treatment of infection or inducing immunological response in

PT mammal  
 PS Claim 10; Page 52-53; 748pp; English.  
 CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein  
 CC sequence of the invention. This sequence was isolated from a HSV-2 strain  
 CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 104.  
 CC Based on homology, this sequence is a (D10879) virion protein  
 CC (Herpes simplex virus).  
 CC The proteins can be used for the treatment or prevention of disease, to  
 CC induce an immunological response in a mammal or to identify inhibitors,  
 CC activators or novel antivirals. Antagonists of the proteins can be used  
 CC to inhibit a viral polypeptide. The DNA sequence or a vector containing  
 CC it can also be used to induce an immunological response in a mammal.  
 SQ Sequence 818 AA;  
 Query Match 83.7%; Score 41; DB 1; Length 818;  
 Best Local Similarity 62.5%; Pred. No. 2.20e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 446 GGGGGDDG 453  
 QY 22 GGGGAAG 29  
 RESULT 4  
 ID W72095 standard; Protein; 1896 AA.  
 AC W72095;  
 DT 18-DEC-1998 (first entry)  
 DE HSV-2 strain SB5 Contig ID 10 ORF#1 protein.  
 KW HSV-2 strain SB5; immunological response induction; therapy;  
 KW antiviral identification; viral protein inhibitor.  
 OS Herpes simplex virus type 2.  
 FH Key Location/Qualifiers  
 FT Misc\_difference 1891..1896  
 FT W09820016-A1.  
 PN 14-MAY-1998.  
 PD 31-OCT-1997; U20016.  
 PR 09-JUN-1997; US-049018.  
 PR 04-NOV-1996; US-030279.  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PI Chan JY, Dabrowski-Anaral CE, Delvecchio AM, Dillon SB,  
 PI Esser KM, Leary JJ;  
 DR WPI: 98-286847/25.  
 DR N-PSDB: V62154.  
 PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention  
 PT and treatment of infection or inducing immunological response in  
 PT mammal  
 PS Claim 10; Page 77-78; 748pp; English.  
 CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein  
 CC sequence of the invention. This sequence was isolated from a HSV-2 strain  
 CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 10.  
 CC Based on homology, this sequence is a large tegument protein.  
 CC The proteins can be used for the treatment or prevention of disease, to  
 CC induce an immunological response in a mammal or to identify inhibitors,  
 CC activators or novel antivirals. Antagonists of the proteins can be used  
 CC to inhibit a viral polypeptide. The DNA sequence or a vector containing  
 CC it can also be used to induce an immunological response in a mammal.  
 SQ Sequence 1896 AA;  
 Query Match 83.7%; Score 41; DB 1; Length 1896;  
 Best Local Similarity 62.5%; Pred. No. 2.20e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 471 GGGGGDDG 478  
 QY 22 GGGGAAG 29  
 RESULT 5  
 ID W72204 standard; Protein; 3119 AA.  
 AC W72204;  
 DT 13-JAN-1999 (first entry)  
 DE HSV-2 strain SB5 Contig ID 15 ORF#39 protein.

KW HSV-2 strain SB5; immunological response induction; therapy;  
 OS antiviral identification; viral protein inhibitor.  
 FN Herpes simplex virus type 2.  
 PN WO9820016-A1.  
 PD 14-MAY-1998.  
 PF 31-OCT-1997; U20016.  
 PR 09-JUN-1997; US-049018.  
 PR 04-NOV-1996; US-030279.  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PI Chan JY, Dabrowski-Amarel CE, Delvecchio AM, Dillon SB,  
 PI Esser KM, Leary JJ;  
 DR WPI: 98-286847/25.  
 DR N-PSDB: V62176.  
 PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention  
 PT and treatment of infection or inducing immunological response in  
 PT mammal  
 PS Claim 10; Page 127-129; 748pp; English.  
 CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein  
 CC sequence of the invention. This sequence was isolated from a HSV-2 strain  
 CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 15.  
 CC Based on homology, this sequence is a large tegument protein.  
 CC The proteins can be used for the treatment or prevention of disease, to  
 CC induce an immunological response in a mammal or to identify inhibitors,  
 CC activators or novel antiviral antagonists of the proteins can be used  
 CC to inhibit a viral polypeptide. The DNA sequence of a vector containing  
 CC it can also be used to induce an immunological response in a mammal.  
 CC Sequence 3119 AA;  
 Query Match 83.7%; Score 41; DB 1; Length 3119;  
 Best Local Similarity 62.5%; Pred. No. 2.20e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 465 GGGGDDG 472  
 QY 22 GGGGAAG 29  
 RESULT 6  
 ID W18651 standard; Protein; 220 AA.  
 AC W18651; 1997 (first entry)  
 DE Human apolipoprotein E gene -1 frameshift mutant product.  
 KW Frameshift mutation product; GAGA motif; somatic mutation; diagnosis;  
 KW detection; antibody; probe; cancer; neoplasia; neurodegenerative;  
 KW Parkinson's; Alzheimer's disease; Pick's; Huntington's disease;  
 KW Down's syndrome; frontal lobe dementia; progressive supranuclear palsy;  
 KW PSP; amyotrophic lateral sclerosis; multiple sclerosis; MS;  
 KW cardiovascular; rheumatoid arthritis.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT misc\_difference 1..220  
 FT /note- "x corresponds to a stop codon in the  
 FT accompanying DNA file, T69792"  
 FT 173..183  
 FT /note- "antigenic peptide used for antibody  
 FT production"  
 FT W09712992-A2.  
 PN 10-APR-1997.  
 PD 02-OCT-1996; IB1105.  
 PE 02-OCT-1995; GB-020080.  
 PR 11-JAN-1996; US-009832.  
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.  
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.  
 PA (UYUT-) UNIV STATE UTRECHT.  
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;  
 PI WPI: 97-226235/20.  
 DR N-PSDB: W18651, W18652.  
 DR Use of mutant genes having frame:shift mutation(s) - for developing  
 PT prods. for the diagnosis, prevention and treatment of associated  
 PT diseases, e.g. cancer or neurodegenerative disease  
 PS Claim 22; Fig 5: 123pp; English.  
 CC W18648 and W18649 are +1 and +2 frameshift mutations, respectively, of  
 CC the human ubiquitin gene nucleotides 1094-1800. This region of the gene

CC contains GAGAG motifs. Frameshift mutants of the tau, ubiquitin,  
 CC apolipoprotein E, microtubule-associated protein 2 (MAP-2), neurofilament  
 CC subunit L, M and H and amyloid A4 genes are claimed. All these genes  
 CC share a common GAGAN motif (N- A, G, C or T), which is the site of common  
 CC GA dinucleotide deletion(s) that cause neurodegenerative disorders.  
 CC Antigenic peptides used for the production of antibodies, and small  
 CC nucleic acid sequences derived from frameshift mutants are used in the  
 CC diagnosis, prevention and treatment of cancer and neurodegenerative  
 CC diseases, e.g. Parkinson's disease, Alzheimer's disease, Down's syndrome,  
 CC frontal lobe dementia (Pick's disease), progressive supranuclear palsy  
 CC (PSP), amyotrophic lateral sclerosis, Huntington's disease, multiple  
 CC sclerosis, and other degenerative diseases such as cardiovascular  
 CC disease and rheumatoid arthritis.  
 CC Sequence 220 AA;  
 Query Match 81.6%; Score 40; DB 1; Length 220;  
 Best Local Similarity 75.0%; Pred. No. 2.83e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Db 38 GGGGASGG'45  
 QY 22 GGGGAAG 29  
 RESULT 7  
 ID R69623 standard; Protein; 381 AA.  
 AC R69623;  
 DE S. avermitilis BCKDH E1-alpha subunit.  
 KW Branched-chain alpha-keto-acid-dehydrogenase; BCKDH; bkd gene;  
 KW avermectin; antibiotic; acaricide; anthelmintic; insecticide;  
 KW nematocide; pesticide.  
 OS Streptomyces avermitilis.  
 PN WO9504150-A.  
 PD 09-FEB-1995.  
 PF 30-MAY-1994; IB0127.  
 PR 30-JUL-1993; US-100518.  
 PA (PFIZ ) PFIZER INC.  
 PI Denova CD;  
 DR WPI: 95-082233/11.  
 DR N-PSDB: Q83793.  
 PT Streptomyces branched-chain alpha-keto-acid dehydrogenase DNA -  
 PT used to improve prodn. of natural avermectin cpds and to produce  
 PT novel avermectin cpds.  
 PS Claim 20; Page 42; 65pp; English.  
 CC Novel branched-chain alpha-keto-acid-dehydrogenase (BCKDH) genes  
 CC (bkd) from Streptomyces avermitilis were cloned using PCR and  
 CC homology probing. DNAs encoding the BCKDH E1-alpha, E1-beta, and  
 CC N-terminal and internal portions of the E2 subunit (Q83793-96)  
 CC were obtained. A genomic sequence including the E1-alpha, E1-beta  
 CC and E2 (partial) bkd ORFs was also isolated (Q83797). Manipulation  
 CC of bkd genes allows the enhanced production of natural or novel  
 CC avermectins.  
 CC Sequence 381 AA;  
 Query Match 79.6%; Score 39; DB 1; Length 381;  
 Best Local Similarity 62.5%; Pred. No. 3.64e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Db 176 GGGGSESG 183  
 QY 22 GGGGAAG 29  
 RESULT 8  
 ID R95636 standard; Protein; 436 AA.  
 AC R95636;  
 DR 25-OCT-1996 (first entry)  
 DE Cartilage-derived morphogenetic protein-2.  
 KW Cattle; cartilage-derived morphogenetic protein-2; CDMP-2;  
 KW articular cartilage; chondrogenic; vulnary; implantation;  
 KW chondromalacia; osteoarthritis; therapy; joint repair.  
 OS Bos taurus.

FH Key Location/Qualifiers
FT region 1. .312
FT modified\_site /note= "Pro-region"
FT 89. .91
FT /note= "N-glycosylation site"
FT cleavage\_site 313. .316
FT /note= "Proteolytic processing site"
FT domain 317. .436
FT /note= "C-terminal mature domain"
FT peptide 352. .382
FT /note= "Consensus conserved motif (R95641)"
PN WO9614335-A1.
PD 17-MAY-1996.
PF 07-NOV-1994; UI2814.
PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA Chang SC, Luyten FP, Moos M;
PI WPI: 96-251714/25.
DR N-PSDB; T31602.
PT New purified cartilage extracts and proteins - used to stimulate the
PT development and repair of cartilage in vivo.
PS Claim 11; Fig 2; 34pp; English.
CC The sequence represents cattle articular cartilage-derived
CC morphogenetic protein-2 (CDMP-2). The N-terminal methionine and
CC signal peptide is missing, but part of the pro-region, a typical
CC proteolytic cleavage site and a C-terminal domain containing 7
CC highly conserved Cys residues characteristic of the transforming
CC growth factor-beta gene family are present. A single N-glycosylation
CC site is located in the pro-region. A consensus highly conserved
CC motif in CDMP proteins (R95641) is present in the C-terminal domain.
CC CDMP-2 is present in a purified cartilage extract (claimed) which
CC stimulates local cartilage formation and repair when combined with a
CC matrix and implanted in a mammal. The protein may be used in therapy
CC of e.g. chondromalacia or osteoarthritis, to heal joint surfaces, or
CC to repair cartilage after reconstructive surgery.
SQ Sequence 436 AA;

Query Match 79.6%; Score 39; DB 1; Length 436;
Best Local Similarity 75.0%; Pred. No. 3.64e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 279 GPGGGAEG 286
QY 22 GGGGAEG 29

RESULT 9
ID R88069 standard; Protein; 464 AA.
AC R88069;
DT 29-OCT-1996 (revised)
DE Human ara Kb beta-galactosidase fusion protein.
KW Atopy-related antigen Kb; ara-2; IgE-autoantigen; atopy;
KW inflammation; diagnosis; therapy; beta-galactosidase fusion protein;
KW glutathione-S-transferase; phage lambda cIT.
OS Homo sapiens.
FH Key Location/Qualifiers
FT peptide 1. .14
FT /note= "beta-galactosidase domain"
FT protein 15. .464
FT /note= "atopy-related antigen Kb"
PN WO9534579-A1.
PD 21-DEC-1995.
PF 14-JUN-1995; SE0723.
PR 14-JUN-1994; SE-002090.
PA (PHAA ) PHARMACIA AB.
PI Kraft D, Natter S, Seiberler S, Valent P, Valenta R;
DR N-PSDB; T09256.
PT Recombinant DNA encoding proteins with atopy-related antigenicity -
PT and related vectors, transformed cells and polypeptide(s), for
PT diagnosis and treatment of inflammation, atopic dermatitis, etc.
PS Claim 7; Page 18-20; 35pp; English.

CC The atopy-related antigen Kb (ara-2) beta-galactosidase fusion
CC protein may be produced recombinantly by vector-mediated
CC beta-galactosidase-, glutathione-S-transferase-, or phage lambda
CC cII-fusion protein gene expression in a prokaryote or eukaryote
CC host cell such as Escherichia coli or human keratinocyte cell line
CC A431. This IgE-autoantigen may be used in the diagnosis of atopy,
CC especially atopic dermatitis or asthma bronchiale, or inflammation,
CC where the antigen is linked to a water-insoluble phase, e.g.
CC microtitre well, strip, dextran, agarose, etc., or is labelled.
CC Alternatively, the autoantigen can be used to measure in vitro
CC cellular reaction against an IgE-autoantigen, and to treat atopy
CC by inducing immunological tolerance.
CC (Entry revised to correct spelling of keyword.)
SQ Sequence 464 AA;

Query Match 79.6%; Score 39; DB 1; Length 464;
Best Local Similarity 71.4%; Pred. No. 3.64e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 289 QGGTADG 295
QY 23 QGGAEG 29

RESULT 10
ID W88442 standard; Protein; 472 AA.
AC W88442;
DT 26-APR-1999 (first entry)
DE Arabidopsis branched chain oxoacid dehydrogenase E1 alpha.
KW Branched chain oxoacid dehydrogenase; BCOADC; transgenic plant;
KW poly(3-hydroxybutyrate-co-3-hydroxyvalerate) copolymer;
KW P(3HB-co-3HV); polyhydroxyalkanoate; biodegradable plastic.
OS Arabidopsis thaliana.
PN WO9900505-A1.
PD 07-JAN-1999.
PF 30-JUN-1998; UI3406.
PR 02-MAR-1998; US-076554.
PR 30-JUN-1997; US-051291.
PR 01-AUG-1997; US-055255.
PR 02-MAR-1998; US-076544.
PA (UMOR ) UNIV MISSOURI.
PI Johnston ML, Luetthy MH, Miernyk JA, Mooney BP, Randall DR;
DR WPI: 99-095750/08.
DR N-PSDB; X06840.
PT New nucleic acid encoding subunits of plastid pyruvate dehydrogenase
PT - or branched chain 2-oxoacid dehydrogenase used for producing
PT polyhydroxyalkanoate polymers in high yield
PS Claim 16; Page 103-104; 151pp; English.
CC This polypeptide comprises the E1 alpha subunit of the
CC branched chain oxoacid dehydrogenase (BCOADC) of Arabidopsis
CC thaliana. The invention provides nucleotide sequences that encode
CC the E1 alpha and E1 beta subunits and E2 component (see X06842-44)
CC of the Arabidopsis BCOADC complex, as well as the E1 alpha and E1
CC beta subunits and E2 component (see X06837-39) of the pyruvate
CC dehydrogenase complex. These nucleotide sequences, and the encoded
CC proteins (see W88439-44), can be introduced into plants in various
CC combinations in order to enhance the conversion of threonine to
CC 2-oxobutylate, propionate, propionyl-CoA, beta-ketovaleryl-CoA and
CC beta-hydroxyvaleryl-CoA. Introduction into such plants of nucleic
CC acid sequences encoding an appropriate beta-keto-thiolase, a
CC beta-ketoacyl-CoA reductase and a polyhydroxyalkanoate synthase
CC will permit such transgenic plants to utilise the increased
CC beta-hydroxyvaleryl-CoA substrate in the production of
CC poly(3-hydroxybutyrate-co-3-hydroxyvalerate) copolymer, a
CC biodegradable plastic used to make mouldings, films, coatings and
CC in drug-release applications.
SQ Sequence 472 AA;

Query Match 79.6%; Score 39; DB 1; Length 472;
Best Local Similarity 62.5%; Pred. No. 3.64e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 264 GDGGTSEG 271



OY 22 GGGAAEG 29  
|:|:|:|

RESULT 11  
ID W64228 standard; Protein: 563 AA.  
AC W64228;  
DT 06-OCT-1998 (first entry)  
DE Human secreted protein from clone C2247\_2.  
KW Secreted protein; human adult testes; nutrition; cytokine; stimulant;  
KW cell proliferation; differentiation; immune system; suppressor; ligand;  
KW regulator; hematopoiesis; tissue growth; activin; inhibitor; haemostatic;  
KW chemotaxis; chemokinetic; thrombosis; receptor; cadherin; tumour;  
KW anti-inflammatory.  
OS Homo sapiens.  
PN WO9827205-A2.  
PD 25-JUN-1998.  
PF 17-DEC-1997; U23330.  
PR 16-DEC-1997; US-991872.  
PR 18-DEC-1996; US-769192.  
PR 13-JAN-1997; US-783401.  
PA (GENY ) GENETICS INST INC.  
PI Agostino MT, Jacobs K, Lavallie ER, McCoy JM, Merberg D,  
PI Racie LA, Spaulding V, Treacy M,  
DR WPI: 98-362774/31.  
DR N-PSDB: V44303.  
DR New polynucleotides and secreted proteins - obtained from human  
PT foetal brain, human adult testes, human adult brain and human adult  
PT salivary gland cDNA libraries  
PS Claim 38j; Page 87-89; 110pp; English.  
CC This sequence represents a novel secreted protein from clone C2247\_2  
CC isolated from a human adult testes cDNA library. This protein has  
CC applications for nutritional use, cytokine and cell  
CC proliferation/differentiation activity, immune stimulating or  
CC suppressing activity, hematopoiesis regulating activity, tissue growth  
CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,  
CC haemostatic and thrombotic activity, receptor/ligand activity,  
CC anti-inflammatory activity, cadherin/tumour invasion suppressor activity,  
CC tumour inhibition activity and other activities.  
SQ Sequence 563 AA;

Query Match 79.6%; Score 39; DB 1; Length 563;  
Best Local Similarity 71.4%; Pred. No. 3.64e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Db 255 QGGTAGD 261  
|:|:|:|  
OY 23 QGGAAEG 29

RESULT 12  
ID W44003 standard; Peptide: 800 AA.  
AC W44003;  
DT 18-JUN-1998 (first entry)  
DE A tumour antigen protein.  
KW Tumour antigen protein; gastric cancer; intracellular digestion;  
KW band; major histocompatibility complex class I antigen; recognition;  
KW T-cell; gene therapy; tumour; autoimmune disease.  
OS Homo sapiens.  
PN WO9746676-A1.  
PD 11-DEC-1997.  
PF 04-JUN-1997; J01893.  
PR 25-NOV-1996; JP-330424.  
PR 07-JUN-1996; JP-168429.  
PR 08-OCT-1996; JP-287572.  
PA (TOH) YTOH K, Shchuljo S;  
PI Inai Y, Itoh K, Shchuljo S;  
DR WPI: 96-042184/04.  
DR N-PSDB: V02016.  
PT DNA encoding tumour antigen protein, fragments of which bind to MHC  
PS class I antigens - useful in gene therapy and auto-immune diseases  
PS Claim 1; Pages 35-39; 49pp; Japanese.  
CC The present sequence represents a tumour antigen protein. It was isolated

CC from a gastric cancer cell line. The tumour antigen protein has the  
CC ability to form fragments by intracellular digestion which bind to major  
CC histocompatibility complex (MHC) class I antigens to form a complex  
CC which is recognised by T-cells. The DNA is useful for gene therapy of  
CC tumours and autoimmune diseases.  
SQ Sequence 800 AA;

Query Match 79.6%; Score 39; DB 1; Length 800;  
Best Local Similarity 71.4%; Pred. No. 3.64e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Db 353 QGGTAGD 359  
|:|:|:|  
OY 23 QGGAAEG 29

RESULT 13  
ID W23331 standard; Protein: 1717 AA.  
AC W23331;  
DT 23-MAR-1998 (first entry)  
DE Neuroblastoma indicative isoform of MAP2 (NB-MAP2).  
KW Microtubule-associated protein 2; MAP2; neuroblastoma; human;  
KW isoform; diagnosis; antibody; PCR primer; probe; NB-MAP2.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Misc\_difference 1711 /note= "encoded by TGA"  
FT Misc\_difference 1716 /note= "encoded by TAG"  
FT  
FT  
PD WO9737323-A1.  
PD 31-JUL-1997.  
PR 23-JAN-1997; E00320.  
PR 23-JAN-1996; EP-100930.  
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
PI Kirsch J, Kurek R;  
DR WPI: 97-393707/36.  
DR N-PSDB: T64877.  
DR Diagnostic agents for detecting neuroblastoma specific isoform of  
PT microtubule-associated protein - for diagnosis or monitoring of  
PT neuroblastoma, are primer-g, probes or antibodies, also new  
PT therapeutic antisense sequence  
PS Disclosure: Page 1: 34pp; English.  
CC This is a 250 KD isoform of the human microtubule-associated protein 2  
CC (MAP2). This isoform (NB-MAP2) is indicative of neuroblastoma. This  
CC isoform NB-MAP2 arises from alternative splicing of the primary MAP2  
CC transcripts. Diagnostic agents for detecting NB-MAP2 in cells can be  
CC provided so as to detect and monitor neuroblastoma. The diagnostic  
CC agents are pairs of primers for specific amplification of DNA  
CC corresponding to at least part of the MAP2 transcript, provided neither  
CC primer binds to a sequence downstream of nucleotide 4168 or upstream of  
CC nucleotide 4510 of the present 5.5 kb sequence, or the corresponding  
CC antisense sequence, an oligonucleotide that binds to NB-MAP2-specific  
CC mRNA or the corresponding DNA in which nucleotides 4168 and 4510 are  
CC next to one another, an antibody specific for an epitope generated by  
CC translation of the region comprising the splice junction of NB-MAP2  
CC specific mRNA, and a cDNA probe obtained by nick translation of random  
CC priming of DNA or RNA specific for NB-MAP and obtainable from tissue or  
CC cell probes. The primers are used to detect MAP2-specific dimorphisms in  
CC tissue and cells by polymerase chain reaction (PCR) analysis. The  
CC oligonucleotide and the probes are used to detect NB-MAP2 specific  
CC transcripts (particularly by binding to membrane-bound mRNA) and the  
CC antibodies are used to detect NB-MAP2 in tissue and cells. Specifically  
CC these tests are used to detect neuroblastoma (or its precursor cells)  
CC and to monitor the success of bone marrow purging. The probes can detect  
CC bone marrow and lung metastases in cases where histopathological  
CC diagnosis of neuroblastoma is difficult. The oligonucleotide is used  
CC therapeutically to inhibit NB-MAP2 expression, resulting in development  
CC of neuroblastoma to normal ganglioneurons  
CC Note: This sequence does not appear in the specification; it has been  
CC made by modifying the human MAP2 sequence provided in Fig 2.  
SQ Sequence 1717 AA;

Query Match 79.6%; Score 39; DB 1; Length 1717;  
Best Local Similarity 71.4%; Pred. No. 3.64e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Db 353 QGGTAGD 359  
|:|:|:|  
OY 23 QGGAAEG 29

Query Match

Best Local Similarity 85.7%; Pred. No. 3.64e+02; Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 34 QGAGGEG 40  
||||:|  
QY 23 QGGAAG 29

RESULT 14  
ID W23329 standard; Protein; 1831 AA.  
AC W23329;  
DT 23-MAR-1998 (first entry)  
DE Microtubule-associated protein 2 (MAP2).  
KW Microtubule-associated protein 2; MAP2; neuroblastoma; human;  
KW isoform; diagnosis; antibody; PCR primer; probe.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Misc\_difference 1825 /note= "encode by TGA"  
FT Misc\_difference 1830 /note= "encoded by TAG"  
FT W09727323-Al.  
PN 31-JUL-1997.  
PD 23-JAN-1997; E00320.  
PR 23-JAN-1996; EP-100930.  
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
PI Kirsch J, Kirsch R;  
DR WPI; 97-393707/36.  
DR N-PSDB; T64850.  
PT Microtubule-associated protein - for diagnosis or monitoring of  
PT neuroblastoma, are primers, probes or antibodies, also new  
PT therapeutic antisense sequence  
PS Disclosure: Fig 2; 34pp; English.  
CC This is the human microtubule-associated protein 2 (MAP2). A 250 kD  
CC isoform of this 270 kD MAP2 is indicative of neuroblastoma. This isoform  
CC NB-MAP2-arises from alternative splicing of the primary MAP2 transcripts.  
CC Diagnostic agents for detecting NB-MAP2 in cells can be provided so as  
CC to detect and monitor neuroblastoma. The diagnostic agents are pairs of  
CC primers for specific amplification of DNA corresponding to at least part  
CC of the MAP2 transcript, provided neither primer binds to a sequence  
CC downstream of nucleotide 4168 or upstream of nucleotide 4510 of the  
CC present 5.5 kb sequence, or the corresponding antisense sequence, an  
CC oligonucleotide that binds to NB-MAP2-specific mRNA or the corresponding  
CC DNA in which nucleotides 4168 and 4510 are next to one another, an  
CC antibody specific for an epitope generated by translation of the region  
CC comprising the splice junction of NB-MAP2 specific mRNA, and a cDNA probe  
CC obtained by nick-translation or random priming of DNA or RNA specific for  
CC NB-MAP and obtainable from tissue or cell probes. The primers are used to  
CC detect MAP2-specific dimorphisms in tissue and cells by polymerase chain  
CC reaction (PCR) analysis. The oligonucleotide and the probes are used to  
CC detect NB-MAP2 specific transcripts (particularly by binding to membrane-  
CC bound mRNA) and the antibodies are used to detect NB-MAP2 in tissue and  
CC cells. Specifically these tests are used to detect neuroblastoma (or its  
CC precursor cells) and to monitor the success of bone marrow purging. The  
CC probes can detect bone marrow and lung metastases in cases where  
CC histopathological diagnosis of neuroblastoma is difficult. The  
CC oligonucleotide is used therapeutically to inhibit NB-MAP2 expression,  
CC resulting in development of neuroblastoma to normal ganglioneurons.  
SQ Sequence 1831 AA;

Query Match 79.6%; Score 39; DB 1; Length 1831;  
Best Local Similarity 85.7%; Pred. No. 3.64e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 34 QGAGGEG 40  
||||:|  
QY 23 QGGAAG 29

RESULT 15  
ID W65053 standard; peptide; 93 AA.  
AC W65053;

10-SEP-1998 (first entry)  
DE E. tenella antigenic protein fragment.  
KW Antigen; immune response; vaccine; pathogen; serology; cross-reactivity;  
KW Apicomplexa; avian coccidiosis; domestic chicken; protection; infection;  
KW broad-spectrum immunity.  
OS Eimeria tenella.  
PN W09818821-Al.  
PD 07-MAY-1998.  
PR 31-OCT-1996; IE00070.  
PR 31-OCT-1996; WO-IE0070.  
PA (FORB-) FORBAIRT.  
PA (FORB-) FORFAS T/A BIORESEARCH IRELAND.  
PI (UVDU-) UNIV COLLEGE DUBLIN.  
PI Mulcahy GVT;  
DR WPI; 98-272136/24.  
PT Vaccine against pathogens inducing cross-reactivity but not  
PT cross-protection - especially parasites from Phylum Apicomplexa.  
PT useful e.g. to immunise chickens against several Eimeria species  
PT causing avian coccidiosis at once  
PS Example 1; Page 16; 32pp; English.  
CC This sequence represents an antigenic protein fragment found in  
CC Eimeria tenella sporozoites and merozoites which induce serological  
CC cross-reactivity but not an accompanying cross-protection within the  
CC Phylum Apicomplexa e.g. Toxoplasma, Cryptosporidium, Babesia, Plasmodium  
CC and especially Eimeria. The genus Eimeria contains at least seven species  
CC causing avian coccidiosis, an economically important disease in domestic  
CC chickens, and the vaccines can confer cross-species protection to  
CC chickens against several Eimeria species. The method involves inducing  
CC broad-spectrum immunity in organisms susceptible to infection by a  
CC relevant pathogen, by selecting a region of an antigen from the pathogen  
CC and binding a series of overlapping sequences from the region to solid  
CC phase supports then probing sequences for reactivity with test sera from  
CC an animal infected with the pathogen and selecting one or more sequences  
CC by failure to bind antibody. The vaccine can provide some cross-species  
CC protection, unlike infection and conventional vaccination, which induce  
CC only species-specific immunity. Since production of oocytes in not  
CC necessary, as in conventional vaccine preparation, the vaccine is more  
CC economical and can be used for broiler chickens as well as more  
CC conventionally vaccinated valuable breeder flocks.  
SQ Sequence 93 AA;

Query Match 77.6%; Score 38; DB 1; Length 93;  
Best Local Similarity 75.0%; Pred. No. 4.66e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 30 GAGGAGGEG 37  
|:|||||  
QY 22 QGGAAGG 29

Search completed: Wed May 10 12:17:49 2000  
Job time : 8 secs.



CC MOLECULE TYPE: protein  
CC DESCRIPTION: predicted primary structure of bovine MARCKS  
CC HYPOTHETICAL: NO  
SQ SEQUENCE 335 AA; 31891 MW; 438716 CN;

Query Match 83.7%; Score 41; DB 2; Length 335;  
Best Local Similarity 75.0%; Pred. No. 1.23e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 178 GEGGAEG 185  
QY 22 GGGGAEG 29

RESULT 2  
ID US-08-726-306A-32 STANDARD; PRT; 77 AA.  
XX  
AC xxxxxx

Sequence 32, Application US/08726306A

Sequence 32, Application US/08726306A  
Patent No. 5958684

GENERAL INFORMATION:

APPLICANT: van Leeuwen, Frederik Willem

APPLICANT: Burbach, Johannes Peter Henri

APPLICANT: Grosveld, Franklin G

TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS

NUMBER OF SEQUENCES: 189

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff, Ltd.

STREET: 1 Financial Center

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02111

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08-726,306A

FILING DATE: 02-Oct-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 95/20080.4

FILING DATE: 02-Oct-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/009,832

FILING DATE: 01-Jan-1996

ATTORNEY/AGENT INFORMATION:

NAME: Williams, Ph.D., Kathleen M.

REGISTRATION NUMBER: 34,380

REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)

TELEPHONE: (617) 345-9100

TELEFAX: (617) 345-9111

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 77 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: unknown

SEQUENCE 77 AA; 7617 MW; 25471 CN;

Query Match 81.6%; Score 40; DB 2; Length 77;  
Best Local Similarity 75.0%; Pred. No. 1.59e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 33 GGGGASG 40

RESULT 4  
ID PCT-US95-02275-24 STANDARD; PRT; 71 AA.  
XX  
AC xxxxxx

Sequence 24, Application PC/TUS9502275

Sequence 24, Application PC/TUS9502275  
GENERAL INFORMATION:

APPLICANT: Wistar Institute of Anatomy & Biology

TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use

TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr, P.O. Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/02275

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/280,443

FILING DATE: 25-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/197,794

FILING DATE: 17-FEB-1994

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.



RESULT 7  
ID US-08-457-459-24 STANDARD; PRT; 71 AA.  
XX  
AC xxxxxx  
DT  
DE  
Sequence 24; Application US/08457459  
XX  
Sequence 24, Application US/08457459  
CC Patent No. 5677428  
CC GENERAL INFORMATION:  
CC APPLICANT: Nishikura, Kazuo  
CC TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use  
CC TITLE OF INVENTION: Thereof  
CC NUMBER OF SEQUENCES: 39  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Howson and Howson  
CC STREET: Spring House  
CC CITY: Spring House  
CC STATE: Pennsylvania  
CC COUNTRY: USA  
CC ZIP: 19477  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION NUMBER: US/08/457,459  
CC FILING DATE: 17-FEB-1994  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/197,794  
CC FILING DATE: 17-FEB-1994  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/280,443  
CC FILING DATE: 25-JUL-1994  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Bak, Mary E.  
CC REGISTRATION NUMBER: 31,215  
CC REFERENCE/DOCKET NUMBER: WST49CUSA  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 215-540-9206  
CC TELEFAX: 215-540-5818  
CC INFORMATION FOR SEQ ID NO: 24:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 71 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
CC SEQUENCE 71 AA; 7734 MW; 21544 CN;  
  
Query Match 79.6%; Score 39; DB 1; Length 71;  
Best Local Similarity 75.0%; Pred. No. 2.04e-02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Db 25 GAGGSAEG 32  
Qy 22 GGGGAAEG 29  
  
RESULT 8  
ID US-08-482-385A-6 STANDARD; PRT; 381 AA.  
XX  
AC xxxxxx  
XX  
DT  
DE  
Sequence 6, Application US/08482385A  
XX  
Sequence 6, Application US/08482385A  
CC Patent No. 5728561  
CC GENERAL INFORMATION:  
CC APPLICANT: DENOYA, CLAUDIO D.  
CC TITLE OF INVENTION: GENES ENCODING BRANCHED CHAIN ALPHA  
CC TITLE OF INVENTION: KETOACID DEHYDROGENASE FROM STREPTOMYCES SPECIES  
CC NUMBER OF SEQUENCES: 15  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: PETER C. RICHARDSON  
CC STREET: 235 EAST 42ND STREET, 20TH FLOOR  
CC CITY: NEW YORK  
CC STATE: NEW YORK  
CC COUNTRY: U.S.A  
CC ZIP: 10017-5755  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/482,385A  
CC FILING DATE: 07-JUN-1995  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: SHEYKA, ROBERT F.  
CC REGISTRATION NUMBER: 31,304  
CC REFERENCE/DOCKET NUMBER: PC8346C  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 212-573-1189  
CC TELEFAX: 212-573-1939  
CC INFORMATION FOR SEQ ID NO: 6:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 381 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC SEQUENCE 381 AA; 41015 MW; 608270 CN;  
  
Query Match 79.6%; Score 39; DB 1; Length 381;  
Best Local Similarity 62.5%; Pred. No. 2.04e-02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
Db 176 GGGGTSEG 183  
Qy 22 GGGGAAEG 29  
  
RESULT 9  
ID US-08-818-024-4 STANDARD; PRT; 539 AA.  
XX  
AC xxxxxx  
XX  
DT  
DE  
Sequence 4, Application US/08818024  
XX  
Sequence 4, Application US/08818024  
CC Patent No. 5965365  
CC GENERAL INFORMATION:  
CC APPLICANT: Bandman, Olga  
CC APPLICANT: Guegler, Karl  
CC APPLICANT: Lal, Preeti  
CC TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN KINASE  
CC NUMBER OF SEQUENCES: 5  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.  
CC STREET: 3174 Porter Drive  
CC CITY: Palo Alto  
CC STATE: CA  
CC COUNTRY: USA  
CC ZIP: 94304  
CC COMPUTER READABLE FORM:

CC Sequence 6, Application US/08482385A  
CC Patent No. 5728561  
CC GENERAL INFORMATION:  
CC APPLICANT: DENOYA, CLAUDIO D.  
CC TITLE OF INVENTION: GENES ENCODING BRANCHED CHAIN ALPHA  
CC TITLE OF INVENTION: KETOACID DEHYDROGENASE FROM STREPTOMYCES SPECIES  
CC NUMBER OF SEQUENCES: 15  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: PETER C. RICHARDSON  
CC STREET: 235 EAST 42ND STREET, 20TH FLOOR  
CC CITY: NEW YORK  
CC STATE: NEW YORK  
CC COUNTRY: U.S.A  
CC ZIP: 10017-5755  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/482,385A  
CC FILING DATE: 07-JUN-1995  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: SHEYKA, ROBERT F.  
CC REGISTRATION NUMBER: 31,304  
CC REFERENCE/DOCKET NUMBER: PC8346C  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 212-573-1189  
CC TELEFAX: 212-573-1939  
CC INFORMATION FOR SEQ ID NO: 6:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 381 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC SEQUENCE 381 AA; 41015 MW; 608270 CN;

Query Match 79.6%; Score 39; DB 1; Length 381;  
Best Local Similarity 62.5%; Pred. No. 2.04e-02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 176 GGGGTSEG 183  
Qy 22 GGGGAAEG 29

RESULT 9  
ID US-08-818-024-4 STANDARD; PRT; 539 AA.  
XX  
AC xxxxxx  
XX  
DT  
DE

Sequence 4, Application US/08818024  
XX  
Sequence 4, Application US/08818024  
CC Patent No. 5965365  
CC GENERAL INFORMATION:  
CC APPLICANT: Bandman, Olga  
CC APPLICANT: Guegler, Karl  
CC APPLICANT: Lal, Preeti  
CC TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN KINASE  
CC NUMBER OF SEQUENCES: 5  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.  
CC STREET: 3174 Porter Drive  
CC CITY: Palo Alto  
CC STATE: CA  
CC COUNTRY: USA  
CC ZIP: 94304  
CC COMPUTER READABLE FORM:

US-08-912-129A-2-09.rai

Thu May 11 06:49:44 2000

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 436  
PRIORITY APPLICATION NUMBER:  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 436  
PRIORITY APPLICATION NUMBER:  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 436  
PRIORITY APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0243 US  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 539 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE: GenBank  
CLONE: 757823  
SEQUENCE 539 AA; 59114 MW; 1430053 CN;

Query Match 79.6%; Score 39; DB 2; Length 539;  
Best Local Similarity 62.5%; Pred. No. 2.04e+02;  
Matches 5; Mismatches 1; Indels 0; Gaps 0;

Db 525 GPGGASDG 532  
QY 22 GQGGAAEG 29

RESULT 10  
ID US-08-912-129A-52 STANDARD; PRT: 373 AA.  
XX  
AC xxxxxx  
DT  
DE  
XX

Sequence 52, Application US/08912129A  
Patent No. 5922533  
GENERAL INFORMATION:  
APPLICANT: VALLARI, ANADRUZELA S.  
APPLICANT: HACKETT, JOHN JR.  
APPLICANT: HICKMAN, ROBERT K.  
APPLICANT: VARITEK, VINCENT A. JR.  
APPLICANT: NECKLAWS, ELIZABETH A.  
APPLICANT: GOLDEN, ALAN M.  
APPLICANT: BRENNAN, CATHERINE A.  
APPLICANT: DEVARE, SUSHIL G.  
TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATION  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS (Windows 95)  
SOFTWARE: Microsoft Word (ASCII format output)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,129A  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 436  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
NAME: Panckers, Andreas M.  
REGISTRATION NUMBER: 32,652  
REFERENCE/DOCKET NUMBER: 6109.US.01  
TELEPHONE: 847-937-9803  
TELEFAX: 847-938-2623  
TELEX:

INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE 373 AA; 42429 MW; 682053 CN;

Query Match 77.6%; Score 38; DB 2; Length 373;  
Best Local Similarity 62.5%; Pred. No. 2.62e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 244 GEGGDEG 251  
QY 22 GQGGAAEG 29

RESULT 11  
ID US-08-912-129A-54 STANDARD; PRT: 618 AA.  
XX  
AC xxxxxx  
DT  
DE  
XX

Sequence 54, Application US/08912129A  
Patent No. 5922533  
GENERAL INFORMATION:  
APPLICANT: VALLARI, ANADRUZELA S.  
APPLICANT: HACKETT, JOHN JR.  
APPLICANT: HICKMAN, ROBERT K.  
APPLICANT: VARITEK, VINCENT A. JR.  
APPLICANT: NECKLAWS, ELIZABETH A.  
APPLICANT: GOLDEN, ALAN M.  
APPLICANT: BRENNAN, CATHERINE A.  
APPLICANT: DEVARE, SUSHIL G.  
TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATION  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS (Windows 95)  
SOFTWARE: Microsoft Word (ASCII format output)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,129A  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 436  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
NAME: Panckers, Andreas M.  
REGISTRATION NUMBER: 32,652  
REFERENCE/DOCKET NUMBER: 6109.US.01  
TELEPHONE: 847-937-9803  
TELEFAX: 847-938-2623  
TELEX:

INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE 373 AA; 42429 MW; 682053 CN;

Query Match 77.6%; Score 38; DB 2; Length 373;  
Best Local Similarity 62.5%; Pred. No. 2.62e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 244 GEGGDEG 251  
QY 22 GQGGAAEG 29

RESULT 11  
ID US-08-912-129A-54 STANDARD; PRT: 618 AA.  
XX  
AC xxxxxx  
DT  
DE  
XX

Sequence 54, Application US/08912129A  
Patent No. 5922533  
GENERAL INFORMATION:  
APPLICANT: VALLARI, ANADRUZELA S.  
APPLICANT: HACKETT, JOHN JR.  
APPLICANT: HICKMAN, ROBERT K.  
APPLICANT: VARITEK, VINCENT A. JR.  
APPLICANT: NECKLAWS, ELIZABETH A.  
APPLICANT: GOLDEN, ALAN M.  
APPLICANT: BRENNAN, CATHERINE A.  
APPLICANT: DEVARE, SUSHIL G.  
TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATION  
NUMBER OF SEQUENCES: 89  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette, 1.44 MB  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS (Windows 95)  
SOFTWARE: Microsoft Word (ASCII format output)





US-091-376-430-2-09.1a1

Search completed: Wed May 10 12:19:04 2000  
Job time : 57 secs.

DB 94 GHGGAPEA 101  
|:|:|:|:  
QV 22 GQGGAEG 29

RESULT 14 STANDARD: 226 AA: PRT:

XX  
AC  
XX  
XX  
DT  
XX  
DE

Patent No. 5187962  
 APPLICANT: SALLHEL, NATASBA V.; BROEKAERT, WILLEM F.; CHUA, NAM-HAI; KUFF, INE  
 INVENTOR: SALLHEL, NATASBA V.; BROEKAERT, WILLEM F.; CHUA, NAM-HAI; KUFF, INE  
 A HEMIN SEQUENCE  
 NUMBER OF SEQUENCES: 7  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/587,071  
 FILING DATE: 24-SEP-1990  
 SEQ ID NO:2:  
 LENGTH: 209  
 SEQUENCE 226 AA: 24141 MW: 284810 CN:

```

Query Match          75.5%; Score 37; DB 4; Length 209;
Best Local Similarity 71.4%; Pred. No. 3.35e+02;
Matches             5; Conservative      2; Mismatches 0; Indels 0; Gaps 0;

```

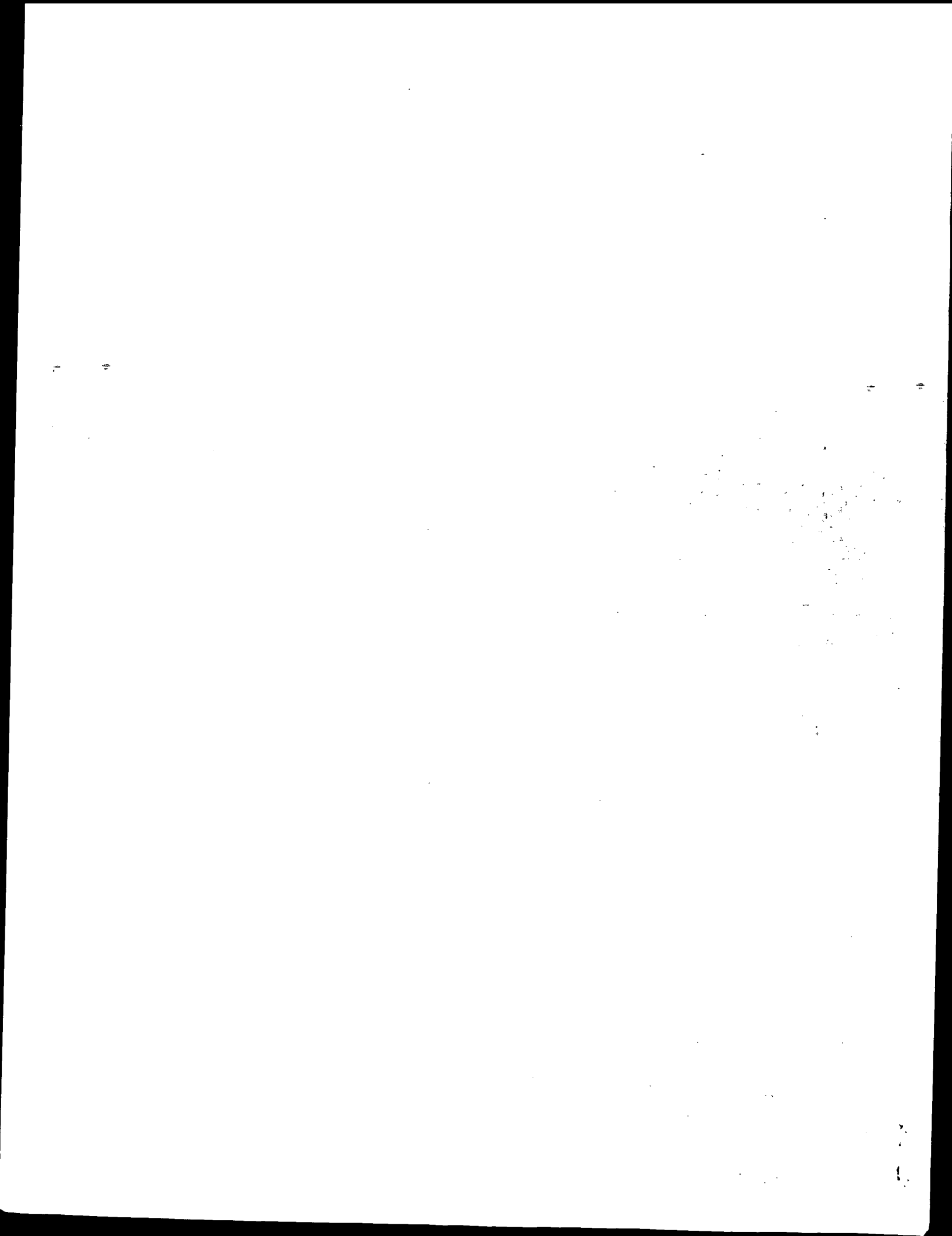
Db	74	QGGSAQ	80
		:	
QY	22	QGGAAE	28

RESULT 15  
TS-08-055-138-6  
STANDARD: 412 AA.  
PRT:

XX AC XX XXXXXX

XX	SEQUENCE 6, Application US/08955138A
DE	
XX	Sequence 6, Application US/08955138A
CC	Patent No. 5977435
CC	GENERAL INFORMATION:
CC	APPLICANT: Lefebvre, Daniel D.
CC	APPLICANT: Gellatly, Kevin S.
CC	TITLE OF INVENTION: PLANT PHOSPHATASES
CC	FILE REFERENCE: PPL97-01
CC	CURRENT APPLICATION NUMBER: US/08/955,138A
CC	CURRENT FILING DATE: 1997-10-21
CC	NUMBER OF SEQ ID NOS: 119
CC	SOFTWARE: FASTSEQ for Windows Version 3.0
CC	SEQ ID NO 6
CC	LENGTH: 412
CC	TYPE: PRT
CC	ORGANISM: SOLANUM TUBerosum
CC	SEQUENCE 412 AA; 47385 MW; 916177 CN;
SO	

	Query Match	75.5%	Score 37	DB 2	Length 412	
	Best Local Similarity	62.5%	Pred. No.	3.35e+02		
	Matches	5	Conservative	2	Mismatches	1
					Indels	0
					Gaps	0
DB	347	GDGENSEG	354			
QY	22	GQGGAAEG	29			





S08341; A32904; S29270; A46098; PS0338  
 S08341  
 #accessions  
 #reference  
 #authors  
 #journal  
 #title  
 #cross-references MUID:89282412  
 #accession  
 #status  
 #molecule\_type mRNA  
 #residues 1-335 #label STU  
 #cross-references EMBL:M24638; NID:gl63339; PID:gl63340  
 A32904  
 #authors  
 #journal  
 #title  
 #cross-references MUID:89282412  
 #accession  
 #status  
 #molecule\_type protein  
 #residues 156-160;165-171;196-215;259-279 #label MIZ  
 #experimental\_source brain  
 #note  
 alanine-rich C-kinase substrate  
 This protein is a major cellular substrate for protein kinase C and plays a role in intracellular signal transduction.  
 It binds to calmodulin in one to one molar ratio in the presence of calcium and cross-links actin filaments.  
 34/3  
 actin binding; blocked amino end; calmodulin binding;  
 lipoprotein; myristylation; phosphoprotein  
 #modified\_site myristylated amino end (Gly) (in mature form) #status predicted

S08341; A32904; S29270; A46098; PS0338  
 S08341  
 #accessions  
 #reference  
 #authors  
 #journal  
 #title  
 #cross-references MUID:89282412  
 #accession  
 #status  
 #molecule\_type mRNA  
 #residues 1-335 #label STU  
 #cross-references EMBL:M24638; NID:gl63339; PID:gl63340  
 A32904  
 #authors  
 #journal  
 #title  
 #cross-references MUID:89282412  
 #accession  
 #status  
 #molecule\_type protein  
 #residues 156-160;165-171;196-215;259-279 #label MIZ  
 #experimental\_source brain  
 #note  
 alanine-rich C-kinase substrate  
 This protein is a major cellular substrate for protein kinase C and plays a role in intracellular signal transduction.  
 It binds to calmodulin in one to one molar ratio in the presence of calcium and cross-links actin filaments.  
 34/3  
 actin binding; blocked amino end; calmodulin binding;  
 lipoprotein; myristylation; phosphoprotein  
 #modified\_site myristylated amino end (Gly) (in mature form) #status predicted

S08341; A32904; S29270; A46098; PS0338  
 S08341  
 #accessions  
 #reference  
 #authors  
 #journal  
 #title  
 #cross-references MUID:89282412  
 #accession  
 #status  
 #molecule\_type mRNA  
 #residues 1-335 #label STU  
 #cross-references EMBL:M24638; NID:gl63339; PID:gl63340  
 A32904  
 #authors  
 #journal  
 #title  
 #cross-references MUID:89282412  
 #accession  
 #status  
 #molecule\_type protein  
 #residues 156-160;165-171;196-215;259-279 #label MIZ  
 #experimental\_source brain  
 #note  
 alanine-rich C-kinase substrate  
 This protein is a major cellular substrate for protein kinase C and plays a role in intracellular signal transduction.  
 It binds to calmodulin in one to one molar ratio in the presence of calcium and cross-links actin filaments.  
 34/3  
 actin binding; blocked amino end; calmodulin binding;  
 lipoprotein; myristylation; phosphoprotein  
 #modified\_site myristylated amino end (Gly) (in mature form) #status predicted

S08341; A32904; S29270; A46098; PS0338  
 S08341  
 #accessions  
 #reference  
 #authors  
 #journal  
 #title  
 #cross-references MUID:89282412  
 #accession  
 #status  
 #molecule\_type mRNA  
 #residues 1-335 #label STU  
 #cross-references EMBL:M24638; NID:gl63339; PID:gl63340  
 A32904  
 #authors  
 #journal  
 #title  
 #cross-references MUID:89282412  
 #accession  
 #status  
 #molecule\_type protein  
 #residues 156-160;165-171;196-215;259-279 #label MIZ  
 #experimental\_source brain  
 #note  
 alanine-rich C-kinase substrate  
 This protein is a major cellular substrate for protein kinase C and plays a role in intracellular signal transduction.  
 It binds to calmodulin in one to one molar ratio in the presence of calcium and cross-links actin filaments.  
 34/3  
 actin binding; blocked amino end; calmodulin binding;  
 lipoprotein; myristylation; phosphoprotein  
 #modified\_site myristylated amino end (Gly) (in mature form) #status predicted

158,162,166,169 #binding\_site phosphate (Ser) (covalent) (by protein kinase C) #status predicted  
 SUMMARY #length 335 #molecular-weight 31979 #checksum 8594  
 Query Match 83.7%; Score 41; DB 2; Length 335;  
 Best Local Similarity 75.0%; Pred. No. 1.93e+01;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Db 178 GEGEAEAG 185  
 QY 22 GQGAAG 29  
 RESULT 3  
 ENTRY  
 TITLE  
 C64230 #type complete  
 pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) El-alpha  
 chain pda - Mycoplasma genitalium (SGC3)  
 #formal\_name Mycoplasma genitalium  
 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change  
 21-Aug-1998  
 C64230  
 A64200  
 #accessions  
 #reference  
 #authors  
 Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.; Bult, C.J.; Kerlavage, A.R.; Sutton, G.; Kelley, J.M.; Fritchman, J.L.; Weidman, J.F.; Small, K.V.; Sandusky, M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.M.; Tomb, J.F.; Dougherty, B.A.; Bott, K.F.; Hu, P.C.; Lucier, T.S.; Peterson, S.N.; Smith, H.O.; Hutchison III, C.A.; Venter, J.C.  
 Science (1995) 270:397-403  
 #journal  
 #title  
 The minimal gene complement of Mycoplasma genitalium.  
 #cross-references MUID:96036346  
 #accession  
 C64230  
 #status  
 preliminary; nucleic acid sequence not shown;  
 translation not shown  
 #molecule\_type DNA  
 #residues 1-358 #label TIGR  
 #cross-references GB:U39706; GB:L43967; NID:gl045962; PID:gl045968;  
 TIGR:MG274  
 #experimental\_source strain G-37  
 GENETICS  
 #genetic\_code SGC3  
 CLASSIFICATION  
 #superfamily pyruvate dehydrogenase (lipoamide) alpha chain;  
 thiamine pyrophosphate-binding domain homology  
 oxidoreductase  
 KEYWORDS  
 FEATURE  
 154-200  
 #domain thiamine pyrophosphate-binding domain homology  
 #label TPB  
 SUMMARY #length 358 #molecular-weight 40651 #checksum 9177  
 Query Match 83.7%; Score 41; DB 2; Length 358;  
 Best Local Similarity 75.0%; Pred. No. 1.93e+01;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Db 163 GDGTAAG 170  
 QY 22 GQGAAG 29  
 RESULT 4  
 ENTRY  
 TITLE  
 S73771 #type complete  
 pyruvate dehydrogenase El-alpha chain - Mycoplasma pneumoniae  
 (ATCC 29342) (SGC3)  
 ALTERNATE\_NAMES  
 #formal\_name Mycoplasma pneumoniae  
 ORGANISM  
 #variety  
 ATCC 29342  
 DATE  
 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change  
 11-Jun-1999  
 #accessions  
 #reference  
 #authors  
 Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.

```
#journal      Nucleic Acids Res. (1996) 24:4420-4449
#title        Complete sequence analysis of the genome of the bacterium
               Mycoplasma pneumoniae.
#cross-references EMBL:AF000044; GB:U00089; NID:g1674130;
#accession     M573771
#status        preliminary; nucleic acid sequence not shown;
               translation not shown
#molecule_type DNA
#residues      1-358 ##label HIM
#cross-references EMBL:AF000044; GB:U00089; NID:g1674130;
               PIDN:AAB6093.1; PID:g1674133
#note          the nucleotide sequence was submitted to the EMBL Data
               Library, November 1996

GENETICS
#gene          pdhA
#genetic_code  SGC3
CLASSIFICATION #superfamily pyruvate dehydrogenase (lipoamide) alpha chain;
               thiamine pyrophosphate-binding domain homology
               #length 358 #molecular-weight 40594 #checksum 7906
SUMMARY
Query Match      83.7%; Score 41; DB 2; Length 358;
Best Local Similarity 75.0%; Pred. No. 1.93e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 163 GGGTAAG 170
Qy 22 GGGTAAG 29

RESULT 5
ENTRY   S59499
TITLE   cellulase egII - smut fungus (Ustilago maydis)
ORGANISM #formal_name Ustilago maydis #common_name corn smut
DATE     20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change
               28-May-1999
ACCESSIONS
REFERENCE S59499
#authors  Schaauwecker, F.; Wanner, G.; Kahmann, R.
#journal  Biol. Chem. Hoppe-Seyler (1995) 376:617-625
#title    Filament-specific expression of a cellulase gene in the
               dimorphic fungus Ustilago maydis.
#cross-references M59499
#accession S59499
#status    preliminary; not compared with conceptual translation
##molecule_type mRNA
##residues 1-393 ##label SCH
#cross-references GB:S81598; NID:g1478378; PIDN:AAB36147.1;
               PID:g1478379
SUMMARY   #length 393 #molecular-weight 39594 #checksum 4355

Query Match      83.7%; Score 41; DB 2; Length 393;
Best Local Similarity 87.5%; Pred. No. 1.93e+01;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 295 GAGGAAG 302
Qy 22 GGGTAAG 29

RESULT 6
ENTRY   S21497
TITLE   ribosomal protein S9 - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE     30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change
               22-Jun-1999
ACCESSIONS
REFERENCE JN0587; FN0499; S21497
#authors  Chan, Y.L.; Paz, V.; Olivera, J.; Wool, I.G.
#journal  Biochem. Biophys. Res. Commun. (1993) 193:106-112
#title    The primary structure of rat ribosomal protein S9.
#cross-references M521497
#accession JN0587
#molecule_type mRNA
```

```
##residues      1-194 ##label CHA
#cross-references EMBL:X66370; NID:g57142; PIDN:CAA47013.1; PID:g57143
#accession     PN0499
#molecule_type protein
#residues      2-28;93-122 ##label CHN
CLASSIFICATION #superfamily rat ribosomal protein S9
KEYWORDS        protein biosynthesis; ribosome
FEATURE         2-194
               #product,ribosomal protein S9 #status predicted #label
               RPS
SUMMARY   #length 194 #molecular-weight 22505 #checksum 9051
Query Match      81.6%; Score 40; DB 1; Length 194;
Best Local Similarity 75.0%; Pred. No. 3.15e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 181 GGGGAGAG 188
Qy 22 GGGGAAG 29

RESULT 7
ENTRY   S55917
TITLE   ribosomal protein S9, cytosolic - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE     28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change
               05-Feb-1999
ACCESSIONS
REFERENCE S55917; S68938
               S55912
#authors  Frigerio, J.M.; Dagorn, J.C.; Iovanna, J.L.
#journal  Biochim. Biophys. Acta (1995) 1262:64-68
#title    Cloning, sequencing and expression of the L5, L21, L27a, L28,
               S5, S9, S10 and S29 human ribosomal protein mRNAs.
#cross-references M55917; S55912; S68938
#accession S55917
#status    nucleic acid sequence not shown
##molecule_type mRNA
##residues 1-194 ##label FRI
#cross-references EMBL:U14966
REFERENCE S68911
#authors  Vladimirov, S.N.; Ivanov, A.V.; Karpova, G.G.; Musolyanov,
               A.K.; Egorov, T.A.; Thiede, B.; Wittmann-Liebold, B.; Otto,
               A.
#journal  Eur. J. Biochem. (1996) 239:144-149
#title    Characterization of the human small-ribosomal-subunit
               proteins by N-terminal and internal sequencing, and mass
               spectrometry.
#cross-references M55917; S68938
#accession S68938
#molecule_type protein
#residues      2-19 ##label VLA
CLASSIFICATION #superfamily rat ribosomal protein S9
KEYWORDS        protein biosynthesis; ribosome
FEATURE         2-194
               #product ribosomal protein S9, cytosolic #status
               experimental #label MAI
SUMMARY   #length 194 #molecular-weight 22571 #checksum 104
Query Match      81.6%; Score 40; DB 2; Length 194;
Best Local Similarity 75.0%; Pred. No. 3.15e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 181 GGGGAGAG 188
Qy 22 GGGGAAG 29

RESULT 8
ENTRY   B70682
TITLE   probable nitrite reductase - Mycobacterium tuberculosis
               (strain H37Rv)
ORGANISM #formal_name Mycobacterium tuberculosis
DATE     17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change
```

```

17-Jul-1998
ACCESSIONS B70682
REFERENCE A70500
#authors Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eigmeier, K.; Gas, S.; Barry III, C.E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
#journal Nature (1998) 393:537-544
#title Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.
#cross-references MUID:98295987
#accession B70682
#status preliminary; nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-563 #label COL
#cross-references GB:281368; GB:AL123456; NID:3261656; PID:e279651; PID:g1655669
#experimental_source strain H37Rv
GENETICS
#gene nira
SUMMARY
#length 563 #molecular-weight 62997 #checksum 1223
Query Match 81.6%; Score 40; DB 2; Length 563;
Best Local Similarity 62.5%; Pred. No. 3.15e+01;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 493 GHGSGVEG 500
|:|:|:|
QY 22 QGGAAEG 29
RESULT 9
ENTRY Tl3618 #type complete
TITLE hypothetical protein EG:808.4 - fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 13-Aug-1999
ACCESSIONS Tl3618
REFERENCE 217695
#authors Papagiannakis, G.; Spanos, L.; Cox, S.; Siden-Kiamos, I.; Louis, C.
#submission submitted to the EMBL Data Library, April 1999
#description Sequencing the distal X chromosome of Drosophila melanogaster.
#accession Tl3618
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-771 #label PAP
#cross-references EMBL:AL022018; NID:e1273253; PID:e1254565; PID:CAA17684.1
GENETICS
#note EG:808.4
SUMMARY
#length 771 #molecular-weight 88529 #checksum 5112
Query Match 81.6%; Score 40; DB 2; Length 771;
Best Local Similarity 75.0%; Pred. No. 3.15e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 725 GEGGAAG 732
|:|:|:|
QY 22 QGGAAEG 29
RESULT 10
ENTRY I50620 #type fragment

```

---

```

prockr2 - chicken (fragment)
#formal_name Gallus gallus #common_name chicken
13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 28-Feb-1997
ACCESSIONS I50620
REFERENCE I50620
#authors Schutz, B.; Niessing, J.
#journal Gene (1994) 148:227-236
#title Cloning and structure of a chicken zinc finger cDNA: restricted expression in developing neural crest cells.
#cross-references MUID:95047430
#accession I50620
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-1173 #label SCH
#cross-references EMBL:X56805; NID:g577018; PID:g577019
GENETICS
#gene ckr2
SUMMARY
#length 1173 #checksum 4476
Query Match 81.6%; Score 40; DB 2; Length 1173;
Best Local Similarity 85.7%; Pred. No. 3.15e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 1113 QGGASEG 1119
|:|:|:|
QY 23 QGGAAEG 29
RESULT 11
ENTRY S42078 #type fragment
TITLE dnaK-type molecular chaperone - white clover (fragment)
ALTERNATE_NAMES heat shock protein 70
ORGANISM #formal_name Trifolium repens #common_name white clover
DATE 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Aug-1999
ACCESSIONS S42078
REFERENCE S42029
#authors Ellison, N.W.
#submission submitted to the EMBL Data Library, February 1994
#description Nucleotide sequence of a cDNA clone encoding a heat shockprotein 70 from white clover (Trifolium repens L.).
#accession S42078
#molecule_type mRNA
#residues 1-127 #label ELL
#cross-references EMBL:Z30176; NID:g456274; PID:CAA82915.1; PID:g456275
FUNCTION
#description involved in protein folding and assembling/disassembling of protein complexes
CLASSIFICATION #superfamily heat shock protein 70
KEYWORDS ATP; molecular chaperone
SUMMARY #length 127 #checksum 1744
Query Match 79.6%; Score 39; DB 2; Length 127;
Best Local Similarity 85.7%; Pred. No. 5.08e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 102 QGGAGEG 108
|:|:|:|
QY 23 QGGAAEG 29
RESULT 12
ENTRY KNRZG2 #type complete
TITLE glycine-rich cell wall structural protein 2 precursor - rice
ORGANISM #formal_name Oryza sativa #common_name rice
DATE 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jul-1999
ACCESSIONS SI8567
REFERENCE SI8567
#authors Fang, R.X.; Pang, Z.; Gao, D.M.; Mang, K.Q.; Chua, N.H.
#journal Plant Mol. Biol. (1991) 17:1255-1257

```

```
#title      cDNA sequence of a virus-inducible, glycine-rich protein gene
#cross-references MUID:92032791
#accession  S18567
#molecule_type mRNA
#residues 1-183 #label FAN
#cross-references EMBL:X4449; NID:920244; PIDN:CAA38315.1; PID:920245
CLASSIFICATION #superfamily glycine-rich cell wall structural protein 1
KEYWORDS      cell wall; structural protein
FEATURE
1-27          #domain signal sequence #status predicted #label SIG\
28-183        #product glycine-rich cell wall structural protein 2
#status predicted #label MAT
SUMMARY        #length 183 #molecular-weight 14920 #checksum 6894

Query Match      79.6%; Score 39; DB 1; Length 183;
Best Local Similarity 75.0%; Pred. NO. 5.08e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 118 GGGGAGG 125
QY 22 GGGGAAEG 29

RESULT 13
ENTRY
TITLE      oleosin - common sunflower
ALTERNATE_NAMES
ORGANISM    oil body protein
DATE        10-Apr-1996 #sequence-revision 19-Apr-1996 #text_change
20-Aug-1999

ACCESSIONS  S60482; S62123; S44070
REFERENCE    S60482
#authors     Thoits, P.J.E.; Millichip, M.I.; Stobart, A.K.; Griffiths,
#journal     W.T.; Shewry, P.R.; Napier, J.A.
#title       Plant Mol. Biol. (1995) 29:403-410
#cross-references MUID:96046759
#accession   S60482
#status      nucleic acid sequence not shown
#molecule_type mRNA
#residues    1-184 #label THO
#cross-references EMBL:X78679

REFERENCE    S62123
#authors     Thoits, P.J.E.; Griffiths, W.T.; Napier, J.A.; Shewry, P.R.
#submission submitted to the EMBL Data Library, December 1995
#accession   S62123
#molecule_type mRNA
#residues    1-63; 'SSDGRAS', 73-184 #label THW
#cross-references EMBL:X78679; NID:9516334; PIDN:CAA55348.1;
PID:91345520

REFERENCE    S44070
#authors     Thoits, P.J.E.; Griffiths, W.T.; Napier, J.A.; Shewry, P.R.
#submission submitted to the EMBL Data Library, April 1994
#accession   S44070
#molecule_type mRNA
#residues    74-126, 'A', 128, 'V', 130, 'R', 132-184 #label THF
#cross-references EMBL:X78679
CLASSIFICATION #superfamily oleosin
KEYWORDS      membrane protein; seed
SUMMARY        #length 184 #molecular-weight 19746 #checksum 5518

Query Match      79.6%; Score 39; DB 2; Length 184;
Best Local Similarity 75.0%; Pred. NO. 5.08e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 171 GGGGKEG 178
QY 22 GGGGAAEG 29

RESULT 14
ENTRY
TITLE      JQ1060 #type fragment
```

```
glycine-rich protein 1 - Arabidopsis thaliana (fragment)
#formal_name Arabidopsis thaliana #common_name mouse-ear
#accession 31-Dec-1991 #sequence-revision 31-Dec-1991 #text_change
20-Aug-1999
#molecule_type mRNA
#residues 1-210 #label DEO
#cross-references GB:S47405; NID:g259442; PIDN:AA24073.1; PID:g259443
#experimental_source strain C24
CLASSIFICATION #superfamily Phaseolus glycine-rich cell wall protein 1.8
SUMMARY        #length 210 #checksum 4387

Query Match      79.6%; Score 39; DB 2; Length 210;
Best Local Similarity 75.0%; Pred. NO. 5.08e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 127 GGGGAGG 134
QY 22 GGGGAAEG 29

RESULT 15
ENTRY
TITLE      B75052 #type complete
coenzyme pqq synthesis protein III related PAB1446 -
#formal_name abyssal (strain Orsay)
#accession 20-Aug-1999 #sequence-revision 20-Aug-1999 #text_change
20-Aug-1999
ACCESSIONS  B75052
REFERENCE    B75052
#authors     anonymous, Genoscope
#submission submitted to the EMBL Data Library, July 1999
#description Pyrococcus abyssal genome sequence: insights into archaeal
chromosome structure and evolution.
#accession   B75052
#status      preliminary
#molecule_type DNA
#residues    1-254 #label KAW
#cross-references GB:AJ248287; GB:AL096836; NID:g5458657;
PIDN:CA50311.1; PID:e1516209; PID:g5458824
#experimental_source strain Orsay

GENETICS
SUMMARY      PAB1446 #molecular-weight 28837 #checksum 8079
#score
#length 254

Query Match      79.6%; Score 39; DB 2; Length 254;
Best Local Similarity 62.5%; Pred. NO. 5.08e+01;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 15 GGGGVEG 22
QY 22 GGGGAAEG 29

Search completed: Wed May 10 12:17:25 2000
Job time : 9 secs.
```









```

CC EMBL; AF000044; AAB96093.1; 1.
DR PFAM; PF00676; EL_dehydrog; 1.
KW Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate;
SQ SEQUENCE 358 AA; 40594 MW; E6184A2026D7A143 CRC64;

Query Match 83.7%; Score 41; DB 1; Length 358;
Best Local Similarity 75.0%; Pred. No. 8.59e+00; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 0;

DB 163 GDGGAAG 170
QY 22 GGGGAAG 29

RESULT 4
ID ODP4_MYCGE STANDARD; PRT; 358 AA.
AC P47516;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE PYRUVATE DEHYDROGENASE E1 COMPONENT, ALPHA SUBUNIT (EC 1.2.4.1).
GN PDHA OR MG274.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE; 96026346.
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky C., Fushman J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.F., Dougherty B.A., Bott K.F., Hu P.C., Luetjens T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL Science 270:397-403 (1995).
CC -1- FUNCTION: THE PYRUVATE DEHYDROGENASE COMPLEX CATALYZES THE OVERALL
CC CONVERSION OF PYRUVATE TO ACETYL-COA & CO(2). IT CONTAINS MULTIPLE
CC COPIES OF THREE ENZYMATIC COMPONENTS: PYRUVATE DEHYDROGENASE (E1),
CC DIHYDROLIPOAMIDE ACETYLTRANSFERASE (E2) & LIPOAMIDE DEHYDROGENASE
CC (E3) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: PYRUVATE + LIPOAMIDE = S-ACETYL-DIHYDRO-
CC LIPOAMIDE + CO(2).
CC -1- COFACTOR: THIAMINE PYRROPHOSPHATE (BY SIMILARITY).
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39707; AAC71496.1; 1.
DR TIGR; MG274; 1.
DR PFAM; PF00676; EL_dehydrog; 1.
KW Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate;
SQ SEQUENCE 358 AA; 40551 MW; 9C85D2335F80842E CRC64;

Query Match 83.7%; Score 41; DB 1; Length 358;
Best Local Similarity 75.0%; Pred. No. 8.59e+00; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 0;

DB 163 GDGGAAG 170
QY 22 GGGGAAG 29

```

```

RESULT 5
ID UNL125MA STANDARD; PRT; 393 AA.
AC P54424;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE ENDOGLUCANASE 1 PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE 1)
DE (CELLULOSE 1) (EG 1).
GN EGL1.
OS Ustilago maydis (Smut fungus).
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PBD11;
RX MEDLINE; 96145728.
RA Schaevecker F., Wanner G., Kahmann R.;
RT "Filament-specific expression of a cellulase gene in the dimorphic
RT fungus Ustilago maydis."
RL Biol. Chem. Hoppe-Seyler 376:617-625 (1995).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC LINKAGES IN CELLULOSE.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: HYPHAL TIP.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN FILAMENTOUS DIKARYON.
CC -1- PTM: MAY ALSO BE O-GLYCOSYLATED.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY K (FAMILY 45 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S81598; AAB36147.1; 1.
DR HSSP; P43316; 2ENG.
DR PROSITE; PS01140; GLYCOSYL-HYDROL_F45; 1.
KW Cellulose degradation; Hydrolase; Glycosidase; Signal; Glycoprotein.
FT SIGNAL 1 26
FT CHAIN 27 393
FT ACT_SITE 34 34 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 152 152 PROTON DONOR (BY SIMILARITY).
FT DOMAIN 270 385 ALA/GLY/SER-RICH.
FT CARBOHYD 343 343 POTENTIAL.
FT SEQUENCE 393 AA; 39594 MW; 65C753C610CD6AD3 CRC64;

Query Match 83.7%; Score 41; DB 1; Length 393;
Best Local Similarity 87.5%; Pred. No. 8.59e+00;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 295 GAGGAAG 302
QY 22 GGGGAAG 29

RESULT 6
ID RS9_RAT STANDARD; PRT; 193 AA.
AC P23314;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE 40S RIBOSOMAL PROTEIN S9.
GN RPS9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-LIVER;
RX MEDLINE; 93277536.

```

RA Chan Y.-L., Paz V., Olvera J., Wool I.G.;  
RT "The primary structure of rat ribosomal protein S9."  
RL Biochem. Biophys. Res. Commun. 193:106-112(1993).  
CC -!- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X66370; CAA47013.1; -  
DR PIR; S21497; S21497.  
DR PROSITE; PS00632; RIBOSOMAL\_S4; 1.  
DR PFAM; PF00163; Ribosomal\_S4; 1.  
DR PFAM; PF01479; S4; 1.  
DR Ribosomal protein.  
FT INIT\_MET 0  
SQ SEQUENCE 193 AA; 22374 MW; BAC3B24AD152E071 CRC64;

Query Match 81.6%; Score 40; DB 1; Length 193;  
Best Local Similarity 75.0%; Pred. No. 1.44e+01;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 180 GGGGAGG 187

||||:|

QY 22 GGGGAAEG 29

RESULT 7

ID RS9 HUMAN STANDARD; PRT; 193 AA.  
AC P46781;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE 40S RIBOSOMAL PROTEIN S9.  
GN RPS9.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eukaryota; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-COLON:  
RX MEDLINE; 95290496.  
RA Prigorio J.-M., Degorn J.C., Iovanna J.L.;  
RT Cloning, sequencing and expression of the L5, L21, L27a, L28, S5,  
RT S9, S10 and S29 human ribosomal protein mRNAs.;  
RL Biochim. Biophys. Acta 1262:64-68(1995).  
CC -!- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U14971; AAA85659.1; -  
DR MIM; 603631; -  
DR PROSITE; PS00632; RIBOSOMAL\_S4; 1.  
DR PFAM; PF00163; Ribosomal\_S4; 1.  
DR PFAM; PF01479; S4; 1.  
DR Ribosomal protein.  
FT INIT\_MET 0  
SQ SEQUENCE 193 AA; 22440 MW; DDA325BFB53384BC CRC64;

Query Match 81.6%; Score 40; DB 1; Length 193;  
Best Local Similarity 75.0%; Pred. No. 1.44e+01;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 180 GGGGAGG 187

||||:|

QY 22 GGGGAAEG 29

RESULT 8

ID PRIO\_CAMDR STANDARD; PRT; 255 AA.  
AC P79141;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE MAJOR PRION PROTEIN PRECURSOR (PRP).  
GN PRP OR PRP  
OS Camelus dromedarius (Dromedary) (Arabian camel).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 98019099.  
RA Kaluz S., Kaluzova M., Flint A.P.F.;  
RT "Sequencing analysis of prion genes from red deer and camel.";  
CC Gene 199:283-286(1997).  
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE  
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.  
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED  
CC "RODS".  
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
CC -!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND  
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,  
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME  
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),  
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.  
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Y09760; CAA70901.1; -  
DR HSP; P04925; IAG2.  
DR PRINTS; PR00341; PRION.  
DR PROSITE; PS00291; PRION\_1; 1.  
DR PROSITE; PS00706; PRION\_2; 1.  
DR PFAM; PF00377; prion; 1.  
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.  
FT SIGNAL 1 24  
FT CHAIN 25 255  
FT DOMAIN 54 94  
FT REPEAT 54 62  
FT REPEAT 63 70  
FT REPEAT 71 78  
FT REPEAT 79 86  
FT REPEAT 87 94  
FT CARBOHYD 184 184  
FT CARBOHYD 200 200  
FT DISULFID 182 217  
SQ SEQUENCE 255 AA; 27595 MW; FABB2DBFA333E494 CRC64;

Query Match 81.6%; Score 40; DB 1; Length 255;  
Best Local Similarity 75.0%; Pred. No. 1.44e+01;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 93 GGGGAGH 100

||||:|

QY 22 GGGGAAEG 29

RESULT 9

Thu May 11 06:49:45 2000

US-09-376-430-2-09.rsp

```

ID VP6_AHSV3          STANDARD;          PRT;    369 AA.
AC Q64909;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VP6 PROTEIN (MINOR INNER CORE PROTEIN VP6).
GN S9.
OS African horse sickness virus 3 (AHSV-3) (African horse sickness virus
(sero type 3)).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96335674.
RA Turnbull P.J., Cormack S.B., Huismans H.;
RT "Characterization of the gene encoding core protein VP6 of two
African horsesickness virus serotypes.";
RL J. Gen. Virol. 77:1421-1423(1996).
CC -1- SIMILARITY: BELONGS TO THE REOVIRUSES VP6 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; U19881; AAB17103.1; -.
DR PFAM; PF01516; Orbi_VP6; 1.
KW Core protein.
SQ SEQUENCE 369 AA; 38464 MW; E596B1EF41A92768 CRC64;

Query Match
Best Local Similarity 81.6%; Score 40; DB 1; Length 369;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 122 GOGAASDG 129
QY 22 GOGGAEG 29

RESULT 10
ID VP6_AHSV6          STANDARD;          PRT;    369 AA.
AC Q64913;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VP6 PROTEIN (MINOR INNER CORE PROTEIN VP6).
GN S9.
OS African horse sickness virus 6 (AHSV-6) (African horse sickness virus
(sero type 6)).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96335674.
RA Turnbull P.J., Cormack S.B., Huismans H.;
RT "Characterization of the gene encoding core protein VP6 of two
African horsesickness virus serotypes.";
RL J. Gen. Virol. 77:1421-1423(1996).
CC -1- SIMILARITY: BELONGS TO THE REOVIRUSES VP6 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; U33000; AAB17107.1; -.
DR PFAM; PF01516; Orbi_VP6; 1.
KW Core protein.
SQ SEQUENCE 369 AA; 38450 MW; 1ECFCC0FFFFF21FE9 CRC64;

```

```

Query Match
Best Local Similarity 81.6%; Score 40; DB 1; Length 369;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 122 GOGAASDG 129
QY 22 GOGGAEG 29

RESULT 11
ID GRP2_ORYSA        STANDARD;          PRT;    183 AA.
AC P29834;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN 2 PRECURSOR.
GN GRP 0.9 OR GRP-1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
OC Poaceae; Oryza.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. A1J1AO NANTE; TISSUE=LEAF;
RX MEDLINE; 92032791.
RA Pang R.X., Pang Z., Gao D.M., Mang K.Q., Chua N.H.;
RT "cDNA sequence of a virus-inducible, glycine-rich protein gene from
rice.";
RL Plant Mol. Biol. 17:1255-1257(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. A1J1AO NANTE INDICA RICE;
RA Pang R.X., Wang Y.L.;
RT Submitted (Nov-1995), to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: CELL WALL (POTENTIAL).
CC -1- INDUCTION: BY INFECTION WITH RICE YELLOW STUNT VIRUS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; X54449; CAA38315.1; -.
DR EMBL; U40708; AAB85863.1; -.
DR PIR; S18567; KNR2G2.
DR HSP; P30129; 4DPV.
KW Cell wall; Structural protein; Repeat; Signal.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 183 GLYCINE-RICH CELL WALL STRUCTURAL
PROTEIN 2.
FT DOMAIN 28 181 GLY-RICH.
SQ SEQUENCE 183 AA; 14920 MW; E9889452101B3893 CRC64;

Query Match
Best Local Similarity 79.6%; Score 39; DB 1; Length 183;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 118 GOGGAEG 125
QY 22 GOGGAEG 29

RESULT 12
ID Y348_MYCE         STANDARD;          PRT;    322 AA.
AC P47590;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)

```

DE HYPOTHETICAL LIPOPROTEIN MG348 PRECURSOR.  
GN MG348.  
OS Mycoplasma genitalium.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
ON Mycoplasmataceae; Mycoplasma.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 33530 / G-37;  
RX MEDLINE; 96026346.  
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
RA Nguyen D.T., Uterberty B.A., Saudek D.M., Phillips C.A., Merrick J.M.,  
RA Tomb J.F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,  
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.,  
RT "The minimal gene complement of Mycoplasma genitalium."  
RL Science 270:397-403(1995).  
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID  
CC ANCHOR (POTENTIAL).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U39716; AAC71573.1; -  
DR TIGR; MG348; -  
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.  
KW Hypothetical protein; Lipoprotein; Membrane; Signal.  
FT SIGNAL 1 27 POTENTIAL.  
FT CHAIN 28 322 HYPOTHETICAL LIPOPROTEIN MG348.  
FT LIPID 28 28 N-ACYL DIGLYCERIDE (POTENTIAL).  
SQ SEQUENCE 322 AA; 35186 MW; BEB2495F470161B3 CRC64;  
Query Match 79.6%; Score 39; DB 1; Length 322;  
Best Local Similarity 50.0%; Pred. No. 2.41e+01;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 257 GEGGSDG 264  
QY 22 GGGGAEG 29  
RESULT 13  
ID PAC1\_SCHPO STANDARD; PRT; 363 AA.  
AC P22192;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE DOUBLE-STRAND-SPECIFIC PAC1 RIBONUCLEASE (EC 3.1.26.3) (HCS PROTEIN).  
GN PAC1 OR HCS OR SPEC119.11C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;  
OC Schizosaccharomycetaceae; Schizosaccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91114707.  
RA Iino Y., Sugimoto A., Yamamoto M.;  
RT "S. pombe pac1+, whose overexpression inhibits sexual development,  
RL encodes a ribonuclease III-like RNase.";  
RN EMO J. 10:221-226(1991).  
DE [2]  
DE VASOPRESSIN VIA RECEPTOR (VASCULAR/HEPATIC-TYPE ARGinine VASOPRESSIN  
DE RECEPTOR) (ANTIDIURETIC HORMONE RECEPTOR 1A) (AVER VIA).  
GN AVPR1A.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Caprinae; Ovis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96004901.

RA MEDLINE; 95342163.  
RX Retondo G., Gillespie M., Frendewey D.;  
RT "Rescue of the fission yeast snRNA synthesis mutant snm1 by  
RT overexpression of the double-strand-specific Pac1 ribonuclease.";  
RL Mol. Gen. Genet. 247:698-708(1995).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972;  
RA Wood V., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.;  
RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP CHARACTERIZATION.  
RX MEDLINE; 96279740.  
RA Retondo G., Frendewey D.;  
RT "Purification and characterization of the Pac1 ribonuclease of  
RT Schizosaccharomyces pombe.";  
RL Nucleic Acids Res. 24:2377-2386(1996).  
CC -!- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. CONVERTS LONG DOUBLE-  
CC STRANDED RNAs INTO SHORT OLIGONUCLEOTIDES, LEAVING 5'-PHOSPHATES  
CC ON THEIR CLEAVAGE PRODUCTS. PROBABLY INHIBITS MATING AND MEIOSIS  
CC BY DEGRADING A SPECIFIC MRNA REQUIRED FOR SEXUAL DEVELOPMENT.  
CC -!- CATALYTIC ACTIVITY: ENDONUCLEOTIC CLEAVAGE TO 5'-PHOSPHO-  
CC MONOESTER.  
CC -!- SIMILARITY: CONTAINS 1 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.  
CC -!- SIMILARITY: BELONGS TO THE RNASE III FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X54998; CAA38745.1; -  
DR EMBL; X53769; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; S78982; AAB34897.1; -  
DR EMBL; AL022117; CAA17926.1; -  
DR PIR; S12605; S12605.  
DR PIR; S13522; S13522.  
DR PROSITE; PS00517; RIBONUCLEASE\_III; 1.  
DR PFAM; PF00636; Ribonuclease\_3; 1.  
DR KW Hydrolase; Nuclease; Endonuclease; RNA-binding; Meiosis.  
FT DOMAIN 338 354 DRBM.  
FT MUTAGEN 178 178 G->D, S: LOSS OF ACTIVITY (MUTANT snm1-1  
FT AND RAN1-114).  
SQ SEQUENCE 363 AA; 41539 MW; BB866CD6AC5AF33A CRC64;  
Query Match 79.6%; Score 39; DB 1; Length 363;  
Best Local Similarity 75.0%; Pred. No. 2.41e+01;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Db 312 GAGGSAEG 319  
QY 22 GGGGAEG 29  
RESULT 14  
ID VIAR\_SHEEP STANDARD; PRT; 418 AA.  
AC P48043;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 13-JUL-1998 (Rel. 36, Last annotation update)  
DE VASOPRESSIN VIA RECEPTOR (VASCULAR/HEPATIC-TYPE ARGinine VASOPRESSIN  
DE RECEPTOR) (ANTIDIURETIC HORMONE RECEPTOR 1A) (AVER VIA).  
GN AVPR1A.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Caprinae; Ovis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96004901.

US-09-376-430-2-09.rsp

Thu May 11 06:49:45 2000

```

RA Hutchins A.-M., Phillips P.A., Venter D.J., Burrell L.M.,
RA Johnston C.I.,
RA "Molecular cloning and sequencing of the gene encoding a sheep
RI arginine vasopressin type 1a receptor".
RI Biochem. Biophys. Acta 1263:266-270(1995).
RL FUNCTION: RECEPTOR FOR ARGinine VASOPRESSIN. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A PHOSPHATIDYL-
CC INOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L41502; AAC41628.1;
CC DR EMBL: L41502; AAC41627.1; ALT_INIT.
CC DR GRCB: GCR 2695;
CC DR PRINTS: PR00237; GPCRHHODOPSN.
CC DR PRINTS: PR00752; VASOPRSNVAR.
CC DR PRINTS: PR00896; VASOPRESSINR.
CC DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
CC DR PFAM: PF00001; 7tm1; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein.
CC FT DOMAIN 1 51
CC FT TRANSMEM 52 75
CC FT DOMAIN 76 87
CC FT TRANSMEM 88 109
CC FT DOMAIN 110 124
CC FT TRANSMEM 125 146
CC FT DOMAIN 147 167
CC FT TRANSMEM 168 189
CC FT DOMAIN 190 217
CC FT TRANSMEM 218 238
CC FT DOMAIN 239 293
CC FT TRANSMEM 294 313
CC FT DOMAIN 314 330
CC FT TRANSMEM 332 351
CC FT DOMAIN 352 418
CC FT DISULFID 123 202
CC FT CARBOHYD 13 13
CC FT CARBOHYD 26 26
CC FT CARBOHYD 195 195
CC FT CARBOHYD 319 319
CC SQ SEQUENCE 418 AA; 46521 MW; BDE60C985A6215CE CRC64;

Query Match 79.6%; Score 39; DB 1; Length 418;
Best Local Similarity 85.7%; Pred.No. 2.41e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 254 QGGPAEG 260
QY 23 QGGAAEG 29

RESULT 15
ID GDF6_BOVIN STANDARD; PRT; 436 AA.
AC P5106;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE GROWTH/DIFFERENTIATION FACTOR 6 PRECURSOR (GDF-6) (CARTILAGE-DERIVED
DE MORPHOGENETIC PROTEIN 2) (CDMP-2) (FRAGMENT).
GN GDF6 OR CDMP2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
OC Bovinae; Bos.
RN [1]

```

```

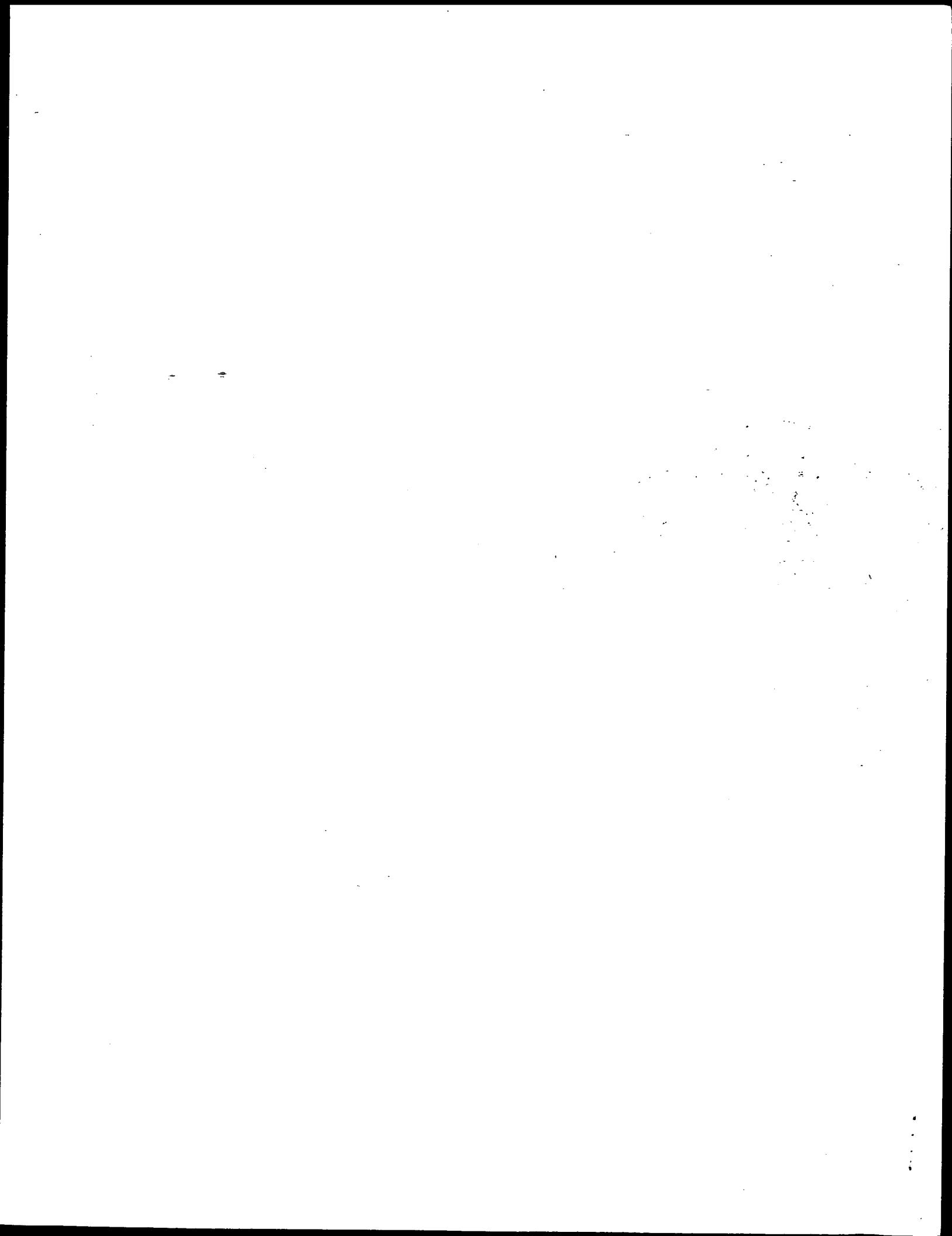
RP SEQUENCE FROM N.A.
RX TISSUE-ARTICULAR CARTILAGE;
RX MEDLINE; 95050604.
RA Chang S., Hoang B., Thomas J.T., Vukicevic S., Luyten F.P.,
RA Ryba N.J.P., Kozak C.A., Reddi A.H., Moos M.;
RT "Cartilage-derived morphogenetic proteins. New members of the
RT transforming growth factor-beta superfamily predominantly expressed
RT in long bones during human embryonic development.";
RL J. Biol. Chem. 269:28227-28234(1994).
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U13661; AAA61416.1;
CC DR PROSITE: PS00250; TGF_BETA; 1.
CC DR PFAM: PF0019; TGF-beta; 1.
CC DR PFAM: PF00688; TGF-beta; 1.
CC KW Growth factor; Cytokine; Glycoprotein.
CC FT NON_TER 1 1
CC FT PROPEP <1 316
CC FT CHAIN 317 436
CC FT DISULFID 335 401
CC FT DISULFID 364 433
CC FT DISULFID 368 435
CC FT DISULFID 400 400
CC FT CARBOHYD 27 27
CC FT CARBOHYD 89 89
CC SQ SEQUENCE 436 AA; 47873 MW; B0688E12EF8AE91D CRC64;

Query Match 79.6%; Score 39; DB 1; Length 436;
Best Local Similarity 75.0%; Pred.No. 2.41e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 279 GGGGAEG 286
QY 22 GGGAAEG 29

Search completed: Wed May 10 12:12:47 2000
Job time : 101 secs.

```







RA COLLINS J.E., HUCKLE E.J.;  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL096880; CAB51404.1; -  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2; 6.  
 DR PROSITE; PS00354; HMG1\_Y; 1  
 KW Hypothetical protein; Zinc-finger; Metal-binding; DNA-binding.  
 SQ SEQUENCE 616 AA; 66319 MW; F7429FB3 CRC32;

Query Match 85.7%; Score 42; DB 4; Length 616;  
 Best Local Similarity 75.0%; Pred. No. 1.05e+01;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 50 GGGGAAG 57  
 QY 22 GGGGAAG 29  
 |:||||:|

RESULT 3  
 ID Q9YER4 PRELIMINARY; PRT; 268 AA.  
 AC Q9YER4;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE 268AA LONG HYPOTHETICAL FKBP-TYPE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE.  
 GN APE0517.  
 OS Aeropyrum pernix.  
 OC Archaea; Crenarchaeota; Aeropyrum.  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=K1;  
 RX MEDLINE; 99310339.

RA KAWARABAYASI Y., HINO Y., HORIKAWA H., YAMAZAKI S., HAIKAWA Y.,  
 RA JIN-NO K., TAKAHASHI M., SEKINE M., BABA S., ANKAI A., KOSUGI H.,  
 RA HOSUYAMA A., FUKUI S., NAGAI Y., NISHIJIMA K., NAKAZAWA H.,  
 RA TAKAMIYA M., MASUDA S., FUNAHASHI T., TANAKA T., KUDOH Y.,  
 RA YAMAZAKI J., KUSHIDA N., OGUCHI A., AOKI K., KUBOTA K., NAKAMURA Y.,  
 RA NOMURA N., SAKO Y., KIKUCHI H.;  
 RT "Complete genome sequence of an aerobic hyper-thermophilic  
 RT Crenarchaeon, Aeropyrum pernix K1.";  
 RL DNA Res. 6:83-101(1999).  
 DR EMBL; AP000059; BAA79482.1; -  
 KW Isomerase.  
 SQ SEQUENCE 268 AA; 29608 MW; BC2E6895 CRC32;

Query Match 83.7%; Score 41; DB 1; Length 268;  
 Best Local Similarity 75.0%; Pred. No. 1.77e+01;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 244 GGGGAAG 251  
 QY 22 GGGGAAG 29  
 |:|||

RESULT 4  
 ID O85699 PRELIMINARY; PRT; 338 AA.  
 AC O85699;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE HYPOTHETICAL 35.5 KD PROTEIN.  
 OS Streptomyces lividans.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=1326, AJ100;  
 RA ALTENBUCHNER J.;  
 RT "Amplifiable element AUD4 from Streptomyces lividans 66.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF072709; AAC25788.1; -  
 DR PFAM; PF01545; Cation\_efflux; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 338 AA; 35529 MW; 92110AEE CRC32;

Query Match 83.7%; Score 41; DB 2; Length 338;  
 Best Local Similarity 75.0%; Pred. No. 1.77e+01;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 164 GGGGLADG 171  
 QY 22 GGGGAAG 29  
 |:||||:|

RESULT 5  
 ID P89459 PRELIMINARY; PRT; 3122 AA.  
 AC P89459;  
 DT 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE VERY LARGE TEGUMENT PROTEIN.  
 GN UL36.  
 OS Herpes simplex virus (type 2).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Simplexvirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HG52;  
 RX MEDLINE; 87111457.  
 RA MCGEOCH D.J., MOSS H.W., MCNAB D., FRAME M.C.;  
 RT "DNA sequence and genetic content of the HindIII 1 region in the short  
 RT unique component of the herpes simplex virus type 2 genome:  
 RT Identification of the gene encoding glycoprotein G, and evolutionary  
 RT comparisons.";  
 RL J. Gen. Virol. 68:19-38(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HG52;  
 RX MEDLINE; 90278430.  
 RA EVERETT R., FENWICK M.;  
 RT "Comparative DNA sequence analysis of the host shutoff genes of  
 RT different strains of herpes simplex virus: type 2 strain HG52 encodes  
 RT a truncated UL41 product.";  
 RL J. Gen. Virol. 71:1387-1390(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HG52;  
 RX MEDLINE; 92113549.  
 RA MCGEOCH D.J., CUNNINGHAM C., MCINTYRE G., DOLAN A.;  
 RT "Comparative sequence analysis of the long repeat regions and  
 RT adjoining parts of the long unique regions in the genomes of herpes  
 RT simplex viruses types 1 and 2.";  
 RL J. Gen. Virol. 72:3057-3075(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HG52;  
 RX MEDLINE; 92356101.  
 RA BARNETT B.C., DOLAN A., TELFORD E.A.R., DAVISON A.J., MCGEOCH D.J.;  
 RT "A novel herpes simplex virus gene (UL49A) encodes a putative membrane  
 RT protein with counterparts in other herpesviruses.";  
 RL J. Gen. Virol. 73:2167-2171(1992).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HG52;  
 RA DOLAN A.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z86099; CAB06723.1; -  
 SQ SEQUENCE 3122 AA; 330047 MW; D2B0A7B8 CRC32;

Query Match 83.7%; Score 41; DB 14; Length 3122;  
 Best Local Similarity 62.5%; Pred. No. 1.77e+01;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 465 GGGGDDG 472  
 QY 22 GGGGAAG 29  
 |:||||:|

Thu May 11 06:49:45 2000

US-09-376-430-2-09.rspt

```

RESULT 6
ID P71753; PRELIMINARY; PRT; 563 AA.
AC 01-FEB-1997 (TREMREL. 02, Created)
DT 01-FEB-1997 (TREMREL. 02, Last sequence update)
DT 01-NOV-1999 (TREMREL. 12, Last annotation update)
DE HYPOTHETICAL 63.0 KD PROTEIN.
GN MTCY253.30C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Mycobacterium.
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA DEVLIN K., CHURCHER C.M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA BARRELL B.G., RAJANDREAM M.A.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
[3]
RN RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA MEDLINE; 96161548.
RA PHILIPP W.J., POULET S., BIGMEIER K., PASCOPELLA L.,
RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA COLE S.T.;
RL "An integrated map of the genome of the tubercle bacillus,
RT Mycobacterium tuberculosis H37RV, and comparison with Mycobacterium
RT leprae."
RL Proc. Natl. Acad. Sci. U.S.A. 93:3132-3137(1996).
DR EMBL; 281368; CAB03734.1; -
DR PFAM; PF01077; NIR-SIR; 1.
DR PRINTS; PR00397; SIROHAEM.
KW Hypothetical protein.
SQ SEQUENCE 563 AA; 62997 MW; 688877C5 CRC32;

Query Match 81.6%; Score 40; DB 2; Length 563;
Best Local Similarity 82.5%; Pred. No. 2.96e+01;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 493 GHGSGVEG 500
Qy 22 GGGAAEG 29

RESULT 7
ID O46307; PRELIMINARY; PRT; 771 AA.
AC 01-JUN-1998 (TREMREL. 06, Created)
DT 01-JUN-1998 (TREMREL. 06, Last sequence update)
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)
DE EG:8D8.4 PROTEIN.
GN EG:8D8.4.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
[1]
RN RP SEQUENCE FROM N.A.
RA PAPAGIANNAKIS G., COX S., SIDEN-KIAMOS I., LOUIS C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN RP SEQUENCE FROM N.A.
RA BENOS P.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL022018; CAA17684.1; -
DR FLYBASE; FBgn0024364; EG:8D8.4.
SQ SEQUENCE 771 AA; 88529 MW; B2D63D51 CRC32;

Query Match 81.6%; Score 40; DB 5; Length 771;

```

```

Best Local Similarity 75.0%; Pred. No. 2.96e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 725 GEGGAANG 732
Qy 22 GGGAAEG 29

RESULT 8
ID Q90783; PRELIMINARY; PRT; 1173 AA.
AC Q90783;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-NOV-1999 (TREMREL. 12, Last annotation update)
DE ZINC FINGER PROTEIN PROCR2 (FRAGMENT).
GN CR2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
[1]
RN RP SEQUENCE FROM N.A.
RC TISSUE-BONE MARROW CELLS;
RX MEDLINE; 95047430.
RA SCHUETZ B., NIESSING J.;
RT "Cloning and structure of a chicken zinc finger cDNA: restricted
RL expression in developing neural crest cells."
RL Gene 148:227-284(1994).
DR EMBL; X56805; CAA40140.1; -
DR HSP; P08047; 1SP2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 15.
DR PFAM; PF00096; zf-C2H2; 16.
KW Zinc finger; Metal-binding; DNA-binding.
FT NON_TER
SQ SEQUENCE 1173 AA; 127327 MW; 2A628B63 CRC32;

Query Match 81.6%; Score 40; DB 13; Length 1173;
Best Local Similarity 85.7%; Pred. No. 2.96e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1113 QGGASEG 1119
Qy 23 GGGAAEG 29

RESULT 9
ID Q41618; PRELIMINARY; PRT; 127 AA.
AC Q41618;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-NOV-1999 (TREMREL. 12, Last annotation update)
DE HEAT SHOCK PROTEIN 70 (FRAGMENT).
OS Trifolium repens (Creeping white clover).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
OC Trifolium.
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN-HUIA; TISSUE-STOLON NODE;
RA ELIISON N.W.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z30176; CAA82915.1; -
DR MENDEL; 15478; Tripp:146;15478.
KW Heat shock.
FT NON_TER
SQ SEQUENCE 127 AA; 14256 MW; F148AD79 CRC32;

Query Match 79.6%; Score 39; DB 10; Length 127;
Best Local Similarity 85.7%; Pred. No. 4.91e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 102 QGGAGEG 108
Qy 1113

```

QY 23 QGGAAEG 29

RESULT 10  
 ID Q39952 AC Q39952; PRELIMINARY; PRT; 183 AA.  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE OLEOSIN.  
 OS Helianthus annuus (Common sunflower).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Asteridae; euasterids II; Asterales; Asteraceae;  
 OC Asteroideae; Heliantheae; Helianthus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SEED EMBRYO;  
 RX MEDLINE; 96046759;  
 RA THOYTS P.J.E., MILLICHIP M.I., STOBART A.K., GRIFFITHS W.T.,  
 RA SHEWRY P.R., NAPIER J.A.;  
 RT "Expression and in vitro targeting of a sunflower oleosin.";  
 RL Plant Mol. Biol. 29:403-410(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SEED EMBRYO;  
 RA THOYTS P.J.E.;  
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X78679; CAA55348.1; -;  
 DR MENDEL; 8425; Helan;1090;8425.  
 DR PFAM; PF01277; Oleosin; 1;  
 SQ SEQUENCE 183 AA; 19719 MW; BA684D3E CRC32;

Query Match 79.6%; Score 39; DB 10; Length 183;  
 Best Local Similarity 75.0%; Pred. No. 4.91e+01;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 170 GGGGKEG 177  
 QY 22 QGGAAEG 29

RESULT 11  
 ID Q41187 AC Q41187; PRELIMINARY; PRT; 210 AA.  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE GLYCINE-RICH PROTEIN (FRAGMENT).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; euroids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE; 93044485.  
 RA DE OLIVEIRA D.E., SEURINCK J., INZE D., VAN MONTAGU M., BOTTERMAN J.;  
 RT "Differential expression of five Arabidopsis genes encoding glycine-  
 RT rich proteins.";  
 RL Plant Cell 2:427-436(1990).  
 DR EMBL; S47405; AAB24073.1; -;  
 DR HSP; P30129; 4DPV.  
 DR MENDEL; 23931; Arath;343;23931.  
 FT NON\_TER 1  
 SQ SEQUENCE 210 AA; 14445 MW; F10138C7 CRC32;

Query Match 79.6%; Score 39; DB 10; Length 210;  
 Best Local Similarity 75.0%; Pred. No. 4.91e+01;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 127 GGGGAGG 134  
 QY 22 QGGAAEG 29

QY 22 QGGAAEG 29

RESULT 12  
 ID Q21588 AC Q21588; PRELIMINARY; PRT; 220 AA.  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)  
 DE M7.10 PROTEIN.  
 GN M7.10.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC LENNARD N.;  
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., PAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SHALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; Z68337; CAA92742.1; -;  
 SQ SEQUENCE 220 AA; 23571 MW; 9B2A53C0 CRC32;

Query Match 79.6%; Score 39; DB 5; Length 220;  
 Best Local Similarity 75.0%; Pred. No. 4.91e+01;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

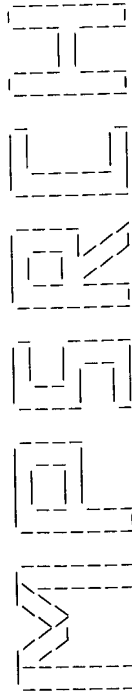
Db 37 GPGGAPEG 44  
 QY 22 QGGAAEG 29

RESULT 13  
 ID Q53080 AC Q53080; PRELIMINARY; PRT; 259 AA.  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE PROTEASOME ALPHA-TYPE SUBUNIT 1.  
 GN PRCA.  
 OS Rhodococcus sp.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NI86/21;  
 RX MEDLINE; 95138028.  
 RA NAGY I., SCHOOF G., COMPERNOLLE F., PROOST P., VANDERLEYDEN J.,  
 RA DE MOT R.;  
 RT "Degradation of the thiocarbamate herbicide EPTC (S-ethyl  
 RT dipropylcarbamothioate) and biosafening by Rhodococcus sp. strain  
 RT NI86/21: involve an inducible cytochrome P-450 system and aldehyde  
 RT dehydrogenase.";  
 RL J. Bacteriol. 177:676-687(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NI86/21;  
 RX MEDLINE; 96082876.  
 RA TAMURA T., NAGY I., LUPAS A., LOTTISPEICH F., CEJKA Z., SCHOOF G.,

RA TANAKA K., DE MOT R., BAUMEISTER W.;  
 RT "The first characterization of a eubacterial proteasome: the 20S  
 RL complex of Rhodococcus";  
 RL Curr. Biol. 5:766-774(1995).  
 DR EMBL; U26421; AAC45741.1; -;  
 KW Proteasome. 259 AA; 28312 MW; DF09825A CRC32;  
 SQ SEQUENCE 259 AA; 28312 MW; DF09825A CRC32;  
 Query Match 79.6%; Score 39; DB 2; Length 259;  
 Best Local Similarity 85.7%; Pred. No. 4.91e+01;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Db 190 GGGAGEG 196  
 |:::|  
 Qy 23 GGGAAEG 29  
 RESULT 14  
 ID Q9XA33 PRELIMINARY; PRT; 261 AA.  
 AC Q9XA33;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE E1-ALPHA BRANCHED-CHAIN ALPHA KETO ACID DEHYDROGENASE (FRAGMENT).  
 GN BKDA2.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA MURPHY L., HARRIS D.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome";  
 RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA BENTLEY S.D., PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE; 97000351.  
 RA REDENBACH M., KIESER H.M., DENAPAITE D., EICHNER A., CULLUM J.,  
 RA KINASHI H., HOPWOOD D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL; AL079308; CAB45198.1; -;  
 RL NON\_TER 1  
 FT NON\_TER 1  
 FT NON\_TER 261  
 SQ SEQUENCE 261 AA; 28003 MW; D80B551F CRC32;  
 Query Match 79.6%; Score 39; DB 2; Length 261;  
 Best Local Similarity 62.5%; Pred. No. 4.91e+01;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Db 176 GGGTSEG 183  
 |:::|  
 Qy 22 GGGAAEG 29  
 RESULT 15  
 ID Q9ZG16 PRELIMINARY; PRT; 336 AA.  
 AC Q9ZG16;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE RNA METHYLTRANSFERASE PIKRI.  
 GN PIKRI.  
 OS Streptomyces venezuelae.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC15439;  
 RX MEDLINE; 98445333.  
 RA XUE Y., ZHAO L., LIU H.W., SHERMAN D.H.;  
 RT "A gene cluster for macrolide antibiotic biosynthesis in Streptomyces  
 RT venezuelae: architecture of metabolic diversity.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:12111-12116(1998).  
 DR EMBL; AF079138; AAC69328.1; -;  
 DR PROSITE; PS01131; RNA\_ALDIMETH; 1.  
 KW Transferase; Methyltransferase.  
 SQ SEQUENCE 336 AA; 37205 MW; C240D98A CRC32;  
 Query Match 79.6%; Score 39; DB 2; Length 336;  
 Best Local Similarity 75.0%; Pred. No. 4.91e+01;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Db 276 GGGGAGG 283  
 |:::|  
 Qy 22 GGGAAEG 29  
 Search completed: Wed May 10 12:16:59 2000  
 Job time : 234 secs.





(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1999-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 12:46:03 2000; Maspar time 227.22 Seconds  
Tabular output not generated. 2.441 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (88-95) from US09376430A.pep (13 of 25)  
Perfect Score: 53  
Sequence: 1 AEQRDDIL 8

Scoring table: PAM 150  
Gap 11

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: streml12  
1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human  
5:sp.invertebrate 6:sp.mammal 7:sp.mhc 8:sp.organelle  
9:sp.phage 10:sp.plant 11:sp.rodent 12:sp.unclassified  
13:sp.vertibrate 14:sp.virus

Statistics: Mean 21.528; Variance 22.723; scale 0.947

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	46	86.8	135	1	027368 HYPOTHETICAL 15.0 KD P	7.38e+00
2	46	86.8	275	2	069053 PTXC.	7.38e+00
3	46	86.8	432	3	060142 HYPOTHETICAL 50.0 KD P	7.38e+00
4	46	86.8	661	14	093130 D8L PROTEIN.	7.38e+00
5	45	84.9	949	4	094766 KIAA0320 PROTEIN (FRAG	1.26e+01
6	45	84.9	3259	4	014789 GIANTIN (GCP372) (MACR	2.14e+01
7	44	83.0	241	2	048861 PLASMIN PSK1 DNA (PAR	2.14e+01
8	44	83.0	271	2	050542 35-KDA PROTEIN.	2.14e+01
9	44	83.0	288	2	008896 REPLICATION PROTEIN RE	2.14e+01
10	44	83.0	293	5	045717 R1A5.2 PROTEIN.	2.14e+01
11	43	81.1	184	14	09WNT4 COAT PROTEIN (FRAGMENT	3.60e+01
12	43	81.1	204	14	09WIP2 NEF.	3.60e+01
13	43	81.1	204	14	09WIP2 NEF.	3.60e+01
14	43	81.1	204	14	09WIP2 NEF.	3.60e+01
15	43	81.1	204	14	09WIP2 NEF.	3.60e+01
16	43	81.1	205	14	09WSE3 NEF PROTEIN.	3.60e+01
17	43	81.1	206	14	09WPU7 NEF PROTEIN.	3.60e+01
18	43	81.1	206	14	09WPU7 NEF PROTEIN.	3.60e+01
19	43	81.1	207	14	09YX73 NEF PROTEIN.	3.60e+01
20	43	81.1	209	14	09WPU5 NEF PROTEIN.	3.60e+01

21	43	81.1	209	14	079778 PROVIRAL NEF.	3.60e+01
22	43	81.1	209	14	079779 PROVIRAL NEF.	3.60e+01
23	43	81.1	209	14	09WPS6 NEF PROTEIN.	3.60e+01
24	43	81.1	211	14	09WPU1 NEF PROTEIN.	3.60e+01
25	43	81.1	211	14	09WPU1 NEF.	3.60e+01
26	43	81.1	212	14	09WPT3 NEF PROTEIN.	3.60e+01
27	43	81.1	213	14	091058 NEF PROTEIN (FRAGMENT)	3.60e+01
28	43	81.1	213	14	091058 NEF PROTEIN.	3.60e+01
29	43	81.1	215	14	09WPU1 COAT PROTEIN.	3.60e+01
30	43	81.1	251	14	071202 COAT PROTEIN.	3.60e+01
31	43	81.1	251	14	041335 309AA LONG HYPOTHETICA	3.60e+01
32	43	81.1	341	2	09VCS8 DIPEPTIDE TRANSPORTER	3.60e+01
33	43	81.1	341	2	056917 TRSE.	3.60e+01
34	43	81.1	452	1	092889 DNA REPAIR PROTEIN RAD	3.60e+01
35	43	81.1	464	2	056755 DNA FOR DHLR, DHLC, DH	3.60e+01
36	43	81.1	468	1	028270 HEME BIOSYNTHESIS PROT	3.60e+01
37	43	81.1	503	5	096044 PAX-2/5/8 (FRAGMENT).	3.60e+01
38	43	81.1	507	5	002016 PAX-B.	3.60e+01
39	43	81.1	703	5	001836 CODED FOR BY C. ELEGAN	3.60e+01
40	43	81.1	772	2	094866 DNA-DEPENDENT RNA POLY	3.60e+01
41	43	81.1	814	11	070162 MYELOID ZINC FINGER PR	3.60e+01
42	43	81.1	955	5	045195 T2282.1 PROTEIN	3.60e+01
43	43	81.1	1252	2	09EJ44 PUTATIVE TYPE II DNA M	3.60e+01
44	43	81.1	3864	5	001808 SIMILARITY TO THE DROS	3.60e+01
45	42	79.2	194	14	09WPU0 NEF PROTEIN.	6.00e+01

ALIGNMENTS

RESULT 1	PRELIMINARY: #	PRT: 135 AA.
ID 027368		
AC 027368		
DI 01-JAN-1998 (TREMBlrel. 05, Created)		
DI 01-JAN-1998 (TREMBlrel. 05, Last sequence update)		
DI 01-AUG-1998 (TREMBlrel. 07, Last annotation update)		
DE HYPOTHETICAL 15.0 KD PROTEIN.		
GN MH313.		
OS Methanobacterium thermoautotrophicum.		
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;		
OC Methanobacterium.		
RN [1]		
RP SEQUENCE FROM N.A.		
RC STRAIN-DELTA H;		
RC MEDLINE; 98037514.		
RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,		
RA ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,		
RA HARRISON D., HOANG L., KEAGLE P., LOMM W., FOHIER S., OIU D.,		
RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,		
RA TIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,		
RA MCDUGALL S., SHIMER G., GOYAL A., PIETROWSKI S., CHURCH G.M.,		
RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;		
RT "Complete genome sequence of Methanobacterium thermoautotrophicum		
RT delta: functional analysis and comparative genomics."		
RL J. Bacteriol. 179:7135-7155(1997).		
DR EMBL; AE000895; AAB85791.1;		
KW Hypothetical protein.		
SQ SEQUENCE 135 AA; 15015 MW; 65CD7501 CRC32;		

Query Match 86.8%; Score 46; DB 1; Length 135;  
Best Local Similarity 85.7%; Pred. No. 7.38e+00;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 45 EHRDDIL 51  
QY 89 EQRDDIL 95

RESULT 2	PRELIMINARY:	PRT: 275 AA.
ID 069053		
AC 069053		
DI 01-AUG-1998 (TREMBlrel. 07, Created)		
DI 01-AUG-1998 (TREMBlrel. 07, Last sequence update)		
DI 01-MAY-1999 (TREMBlrel. 10, Last annotation update)		

DE PTXC.  
 GN PTXC.  
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;  
 OC Pseudomonas.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WM88;  
 RA METCALF W.W., WOLFE R.S.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF061070; AAC71708.1; -;  
 DR PFAM; PF00528; BPD\_transp; 1;  
 SQ SEQUENCE 275 AA; 29396 MW; 7CFC2969 CRC32;

Query Match 86.8%; Score 46; DB 2; Length 275;  
 Best Local Similarity 75.0%; Pred. No. 7.38e+00;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 12 AEQREHIL 19

QY 88 AEQRDDIL 95

RESULT 3  
 ID O60142 PRELIMINARY; PRT; 432 AA.  
 AC O60142;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
 DE HYPOTHETICAL 50.0 KD PROTEIN.  
 GN SPBC18H10.11C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972H-;  
 RA LYNE M., RAJANDREAM M.A., BARRELL B.G., BADCOCK K., CHURCHER C.M.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL022304; CAA18408.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 432 AA; 50048 MW; FE1399FD CRC32;

Query Match 86.8%; Score 46; DB 3; Length 432;  
 Best Local Similarity 50.0%; Pred. No. 7.38e+00;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 241 PEOREEVL 248

QY 88 AEQRDDIL 95

RESULT 4  
 ID O93130 PRELIMINARY; PRT; 661 AA.  
 AC O93130;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE DBL PROTEIN.  
 GN DBL.  
 OS Cowpox virus (CPV).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Orthopoxvirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GRI-90;  
 RX MEDLINE; 98229462.  
 RA SCHCHUNKOV S.N., SAFRONOV P.F., TOTMENIN A.V., PETROV N.A.,  
 RA RYAZANKINA O.I., GUTOROV V.V., KOTVAL G.J.;  
 RT The genomic sequence analysis of the left and right species-specific  
 RT terminal region of a cowpox virus strain reveals unique sequences and  
 RT a cluster of intact ORFs for immunomodulatory and host range

RT proteins.";  
 RL Virology 243:432-460(1998).  
 DR EMBL; Y11842; CAA72584.1; -;  
 DR PFAM; PF00023; ank; 1;  
 SQ SEQUENCE 661 AA; 78051 MW; D78F5A1A CRC32;

Query Match 86.8%; Score 46; DB 14; Length 661;  
 Best Local Similarity 85.7%; Pred. No. 7.38e+00;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 471 QORDDIL 477

QY 89 EQRDDIL 95

RESULT 5  
 ID O9Y4G6 PRELIMINARY; PRT; 949 AA.  
 AC O9Y4G6;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE KIAA0320 PROTEIN (FRAGMENT).  
 GN KIAA0320.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE; 97349984.  
 RA NAGASE T., ISHIKAWA K., NAKAJIMA D., OHIRA M., SEKI N., MIYAJIMA N.,  
 RA TANAKA A., KOFANI H., NOMURA N., OHARA O.;  
 RT "Prediction of the coding sequences of unidentified human genes. VII.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro."  
 RL DNA Res. 4:141-150(1997).  
 DR EMBL; AB002318; BAA20778.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 949 AA; 100503 MW; 04E9FE9A CRC32;

Query Match 84.9%; Score 45; DB 4; Length 949;  
 Best Local Similarity 50.0%; Pred. No. 1.26e+01;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 421 ADHRENIL 428

QY 88 AEQRDDIL 95

RESULT 6  
 ID Q14789 PRELIMINARY; PRT; 3259 AA.  
 AC Q14789; O14398;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE GRANTIN (GCP372) (MACROGOLGIN) (GOLGI AUTOANTIGEN, GOLGIN SUBFAMILY B,  
 DE 1).  
 GN GOLGBL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94187728;  
 RA SEELIG H.P., SCHRANZ P., SCHROETER H., WIEMANN C., GRIFFITHS G.,  
 RA RENZ M.;  
 RT "Molecular genetic analyses of a 376-kilodalton Golgi complex membrane  
 RT protein (grantin) [retracted in Mol Cell Biol 1995 Jan;15(1):591].";  
 RL Mol. Cell. Biol. 14:2564-2576(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94257116.  
 RA SEELIG H.P., SCHRANZ P., SCHROETER H., WIEMANN C., GRIFFITHS G.,



US-09-376-430-2-13.rspt

Thu May 11 06:49:54 2000

RENZ M.;  
 "Macrogolgin--a new 376 kD Golgi complex outer membrane protein as  
 target of antibodies in patients with rheumatic diseases and HIV  
 infections";  
 J. Autoimmun. 7:67-91(1994).  
 [3]  
 SEQUENCE FROM N.A.  
 RP MEDLINE; 95100974.  
 RA SHODA M., MISUMI Y., FUJIWARA T., NISHIOKA M., IKEHARA Y.;  
 "Molecular cloning and sequence analysis of a human 372-kDa protein  
 localized in the Golgi complex";  
 Biochem Biophys Res Commun. 205:1399-1408(1994).  
 CC -|- FUNCTION: MAY PARTICIPATE IN FORMING INTERCISTERNAL CROSS-BRIDGES  
 OF THE GOLGI COMPLEX.  
 CC -|- SUBUNIT: DISULFIDE-LINKED HOMODIMER.  
 CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. GOLGI MEMBRANE.  
 CC -|- ALTERNATIVE PRODUCTS: DIFFERENT ISOFORMS MAY ARISE BY ALTERNATIVE  
 SPLICING. ANTIGEN IN CHRONIC RHEUMATOID ARTHRITIS AND IN THE  
 AUTOIMMUNE DISEASE SJOEGREN'S SYNDROME.  
 CC -|- DISEASE: ANTIGEN IN CHRONIC RHEUMATOID ARTHRITIS AND IN THE  
 AUTOIMMUNE DISEASE SJOEGREN'S SYNDROME.  
 DR EMBL; X75304; CAA53052.1; -.  
 DR EMBL; D25542; BAA05025.1; -.  
 DR MM; 602500; -.  
 KW Golgi stack; Antigen; Coiled coil; Transmembrane;  
 Alternative splicing.  
 FT DOMAIN 1 3235  
 FT TRANSMEM 3236 3256  
 FT DOMAIN 3257 3259  
 FT DOMAIN 48 110  
 FT DOMAIN 127 223  
 FT DOMAIN 238 448  
 FT DOMAIN 460 526  
 FT DOMAIN 545 593  
 FT DOMAIN 677 956  
 FT DOMAIN 969 1028  
 FT DOMAIN 1062 1128  
 FT DOMAIN 1154 1245  
 FT DOMAIN 1301 1779  
 FT DOMAIN 1828 2781  
 FT DOMAIN 2797 2857  
 FT DOMAIN 2872 2993  
 FT DOMAIN 3026 3102  
 FT DOMAIN 3133 3195  
 FT DOMAIN 3420 3423  
 FT DOMAIN 2993 2996  
 FT VARSPLIC 1 39  
 FT VARSPLIC 215 215  
 FT CONFLICT 1765 1765  
 FT CONFLICT 2950 2950  
 FT CONFLICT 2950 2950  
 SQ SEQUENCE 3259 AA; 376072 MW; D947BE6B CRC32;  
 Query Match 84.9%; Score 45; DB 4; Length 3259;  
 Best Local Similarity 75.0%; Pred. No. 1.26e+01;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Db 285 AEORNOIL 292  
 QY 88 AEORDDIL 95  
 RESULT 7  
 ID Q48861 PRELIMINARY; PRT; 241 AA.  
 AC Q48861; 1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE PLASMID PSA1 DNA (PARTIAL).  
 OS Lactobacillus sake.  
 OG Plasmid PSA1.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
 OC Lactobacillus.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=LTH679;  
 RA KLEIN J.R., HENRICH B., ULRICH C.;  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; 250862; CAA90734.1; -.  
 KW Plasmid  
 SQ SEQUENCE 241 AA; 28424 MW; E09E8F13 CRC32;  
 Query Match 83.0%; Score 44; DB 2; Length 241;  
 Best Local Similarity 50.0%; Pred. No. 2.14e+01;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Db 7 PDREDIL 14  
 QY 88 AEORDDIL 95  
 RESULT 8  
 ID Q50542 PRELIMINARY; PRT; 271 AA.  
 AC Q50542;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE 35-KDA PROTEIN.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RA;  
 RX MEDLINE; 90378140.  
 RA O'CONNOR S.P., RUMSCHLAG H.S., MAYER L.W.;  
 "Nucleotide sequence of the gene encoding the 35-kDa protein of  
 Mycobacterium tuberculosis";  
 Res Microbiol. 141:407-423(1990).  
 RL EMBL; M69187; AAA73064.1; -.  
 DR EMBL; M69187; AAA73064.1; -.  
 SQ SEQUENCE 271 AA; 29675 MW; 650F432C CRC32;  
 Query Match 83.0%; Score 44; DB 2; Length 271;  
 Best Local Similarity 62.5%; Pred. No. 2.14e+01;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 212 AEORHEVL 219  
 QY 88 AEORDDIL 95  
 RESULT 9  
 ID Q08896 PRELIMINARY; PRT; 288 AA.  
 AC Q08896;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-FEB-1997 (TREMBLrel. 02, Last annotation update)  
 DE REPLICATION PROTEIN REPA.  
 GN REPA.  
 OS Escherichia coli.  
 OG Plasmid PRAL.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HMS50;  
 RA LLANES C., GABANT P., COUTURIER M., MICHEL-BRIAND Y.;  
 J. Bacteriol. 76:3403-3407(1993).  
 CC -|- FUNCTION: REPA IS ESSENTIAL FOR ORIGIN FUNCTION. AUTOREGULATES ITS  
 OWN SYNTHESIS FROM THE PROMOTER AND WHEN OVERPRODUCED, BLOCKS  
 ORIGIN FUNCTION.  
 CC EMBL; X73674; CAA52023.1; -.  
 DR Plasmid; DNA replication; DNA-binding.  
 KW Plasmid; DNA replication; DNA-binding.  
 SQ SEQUENCE 288 AA; 32992 MW; 69533BDD CRC32;  
 Query Match 83.0%; Score 44; DB 2; Length 288;  
 Best Local Similarity 62.5%; Pred. No. 2.14e+01;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 110 AGQREIL 117  
| : : : |  
QY 88 AEQRDDIL 95

RESULT 10  
ID Q45717 PRELIMINARY; PRT; 293 AA.  
AC Q45717;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)  
DE R1A5.2 PROTEIN.  
GN R1A5.2  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MCMURRAY A.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA WILSON R., AINSKOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY I., COOPER J., COULSON A.,  
RA CRAYTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIRRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans".  
RL Nature 368:32-38(1994).  
DR EMBL; 281122; CAB05597.1; -.  
SQ SEQUENCE 293 AA; 34365 MW; CEE2DC34 CRC32;

Query Match 83.0%; Score 44; DB 5; Length 293;  
Best Local Similarity 71.4%; Pred. No. 2.14e+01;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 109 EQRNDVL 115  
| : : : |  
QY 89 EQRDDIL 95

RESULT 11  
ID Q9WNT4 PRELIMINARY; PRT; 184 AA.  
AC Q9WNT4;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
DE COAT PROTEIN (FRAGMENT).  
GN AVL.  
OS tomato yellow mosaic virus.  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VENEZUELA;  
RA MORALES F.J., LASTRA J.R., DE ULCATEGUI R.C., CALVERT L.A.;  
RT "Potato yellow mosaic geminivirus: a synonym of tomato yellow mosaic  
begomovirus."  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF150742; AAD43532.1; -.  
KW Coat protein.  
FT NON\_TER 1  
FT NON\_TER 184  
SQ SEQUENCE 184 AA; 21513 MW; 3B25E73F CRC32;

Query Match 81.1%; Score 43; DB 14; Length 184;

Best Local Similarity 85.7%; Pred. No. 3.60e+01;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 26 EQRHDL 32  
| : : : |  
QY 89 EQRDDIL 95

RESULT 12  
ID Q9WIP2 PRELIMINARY; PRT; 204 AA.  
AC Q9WIP2;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
DE NEF.  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviridae; Lentivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SUBJECT S3;  
RA MCNEARNEY T., HORNICKOVA Z., TEMPLETON A., BIRDWELL A., ARENS M.,  
RA MARKHAM R., SAAH A., RATNER L.;  
RT "Nef and LTR Sequence Variation from Sequentially Derived Human  
Immunodeficiency Virus Type 1 Isolates."  
RL Virology 208:388-398(1995).  
DR EMBL; U03338; AAA58286.1; -.  
SQ SEQUENCE 204 AA; 23242 MW; 6EAFE028 CRC32;

Query Match 81.1%; Score 43; DB 14; Length 204;  
Best Local Similarity 62.5%; Pred. No. 3.60e+01;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 102 SQQRDIL 109  
| : : : |  
QY 88 AEQRDDIL 95

RESULT 13  
ID Q9WIP3 PRELIMINARY; PRT; 204 AA.  
AC Q9WIP3;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
DE NEF.  
GN NEF.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviridae; Lentivirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SUBJECT S3;  
RA MCNEARNEY T., HORNICKOVA Z., TEMPLETON A., BIRDWELL A., ARENS M.,  
RA MARKHAM R., SAAH A., RATNER L.;  
RT "Nef and LTR Sequence Variation from Sequentially Derived Human  
Immunodeficiency Virus Type 1 Isolates."  
RL Virology 208:388-398(1995).  
DR EMBL; U03340; AAA58288.1; -.  
SQ SEQUENCE 204 AA; 23173 MW; D2C96005 CRC32;

Query Match 81.1%; Score 43; DB 14; Length 204;  
Best Local Similarity 62.5%; Pred. No. 3.60e+01;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 102 SQQRDIL 109  
| : : : |  
QY 88 AEQRDDIL 95

RESULT 14  
ID Q79781 PRELIMINARY; PRT; 204 AA.  
AC Q79781;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

US-09-376-430-2-13.rspt

Thu May 11 06:49:54 2000

01-NOV-1999 (TREMELrel. 12, Last annotation update)  
 DT PROVIRAL NEF.  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroviridae; Lentivirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=12;  
 EX MEDLINE; 93323205.  
 RA SHUGARS D.C.; SMITH M.S.; GLUECK D.H.; NANTERMET P.V.,  
 RA SELLIER-MOISEWITSCH F.; SWANSTROM R.;  
 RT "Analysis of human immunodeficiency virus type 1 nef gene sequences  
 present in vivo.";  
 RL J. Virol. 67:4639-4650(1993).  
 DR EMBL; L15508; AAA02662.1; -;  
 DR HSP; P03406; IEFN.  
 DR PFAM; PF00469; F-protein; 1.  
 SQ SEQUENCE 204 AA; 23459 MW; 3452A151 CRC32;  
 Query Match 81.1%; Score 43; DB 14; Length 204;  
 Best Local Similarity 62.5%; Pred. No. 3.60e+01;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Db 101 SQQRDIL 108  
 QY 88 AEQRDIL 95  
 RESULT 15 PRELIMINARY; PRT; 204 AA.  
 AC Q79780;  
 DT 01-NOV-1996 (TREMELrel. 01, Created)  
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
 DT 01-NOV-1999 (TREMELrel. 12, Last annotation update)  
 DE PROVIRAL NEF.  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroviridae; Lentivirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=13;  
 EX MEDLINE; 93323205.  
 RA SHUGARS D.C.; SMITH M.S.; GLUECK D.H.; NANTERMET P.V.,  
 RA SELLIER-MOISEWITSCH F.; SWANSTROM R.;  
 RT "Analysis of human immunodeficiency virus type 1 nef gene sequences  
 present in vivo.";  
 RL J. Virol. 67:4639-4650(1993).  
 DR EMBL; L15509; AAA02663.1; -;  
 DR HSP; P03406; IEFN.  
 DR PFAM; PF00469; F-protein; 1.  
 SQ SEQUENCE 204 AA; 23464 MW; 2012CFAB CRC32;  
 Query Match 81.1%; Score 43; DB 14; Length 204;  
 Best Local Similarity 62.5%; Pred. No. 3.60e+01;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Db 101 SQQRDIL 108  
 QY 88 AEQRDIL 95

Search completed: Wed May 10 12:50:00 2000  
 Job time : 237 secs.



\*\*\*\*\*  
 W P S R L H  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 12:58:55 2000; MasPar time 3.10 Seconds  
 Tabular output not generated. 53.472 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (99-105) from US09376430A.pap (14 of 25)  
 Perfect Score: 50  
 Sequence: 1 IRNGTHP 7

Scoring table: PAM 150  
 Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: a-geneseq35  
 1:geneseqp

Statistics: Mean 14.061; Variance 31.299; scale 0.449

pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	43	86.0	833	1 W94058	Murine MSH5 (mMSH5) pr	3.43e+01
2	41	82.0	1957	1 W21737	Wild type rat DRG (SNS	6.65e+01
3	41	82.0	1957	1 W21740	Variant rat DRG (SNS-B	6.65e+01
4	41	82.0	2132	1 W21739	Variant rat DRG (SNS-B	6.65e+01
5	40	80.0	258	1 W08137	Human cytokine respons	9.22e+01
6	40	80.0	258	1 W38319	Human cytokine-inducib	9.22e+01
7	40	80.0	415	1 R55197	Rabbit zona pellucida	9.22e+01
8	40	80.0	415	1 W81807	Rabbit ZPC protein.	9.22e+01
9	38	76.0	135	1 R41945	Serotonin receptor 5-H	1.76e+02
10	38	76.0	479	1 R41942	Serotonin receptor 5-H	1.76e+02
11	38	76.0	481	1 W40814	Human 5-HT2B receptor	1.76e+02
12	38	76.0	481	1 R71033	G protein protein coup	1.76e+02
13	38	76.0	504	1 R54682	Mouse brain 5HT2C sero	1.76e+02
14	38	76.0	550	1 W81351	Human guanine nucleoti	1.76e+02
15	38	76.0	580	1 W81349	Human guanine nucleoti	1.76e+02
16	38	76.0	623	1 W94308	Coriolus versicolor py	1.76e+02
17	38	76.0	623	1 R99628	Pyranose oxidase.	1.76e+02
18	38	76.0	941	1 W68203	M. catarrhalis strain	1.76e+02
19	37	74.0	344	1 Y02670	Human secreted protein	2.41e+02
20	37	74.0	1134	1 W46859	Bacillus thuringiensis	2.41e+02
21	37	74.0	1156	1 W46857	Bacillus thuringiensis	2.41e+02
22	37	74.0	1157	1 W84567	Amino acid sequence of	2.41e+02
23	37	74.0	1157	1 W84581	Amino acid sequence of	2.41e+02

## ALIGNMENTS

RESULT 1  
 ID W94058 standard; Protein; 833 AA.  
 AC W94058.  
 DT 06-APR-1999 (first entry)  
 DE Murine MSH5 (mMSH5) protein.  
 KW MSH5; mMSH5; DNA mismatch repair gene; chromosome segregation; meiosis;  
 KW malignant; infertility; Down's syndrome; tumour; cancer; gene therapy;  
 OS murine.  
 PN W09901350-A1.  
 PD 14-JAN-1999.  
 PR 02-JUL-1998; US-051686.  
 PR 03-JUL-1997; US-051686.  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 PI Kolodner R, Winand N;  
 DR WPI; 99-106052/09.  
 DR N-PSDB; X05163.  
 PT New isolated human DNA mismatch repair gene, MSH5 - used for  
 PT developing products for the diagnosis and therapy of disorders such  
 PT as cancer, infertility and Down's syndrome  
 PS Disclosure; Page 61; 114pp; English.  
 CC The invention relates to a human DNA mismatch repair gene, MSH5. Host  
 CC cells containing a vector comprising the MSH5 gene is used for the  
 CC recombinant production of the MSH5 protein. The MSH5 gene product is  
 CC required for meiotic crossing over and segregation of chromosomes during  
 CC meiosis. The products can be used for detecting an alteration in a  
 CC mammalian gene as indicative of a predisposition to a malady associated with  
 CC cells or indicative of a predisposition to a malady associated with  
 CC inappropriate meiotic segregation such as infertility or Down's syndrome.  
 CC The alterations can also be used for diagnosing a DNA mismatch pair  
 CC defective tumour and for prognosis of an individual having cancer.  
 CC Moreover, defects in this gene confer resistance to alkylating agents.  
 CC The products can also be used to identify therapeutic agents effective  
 CC against MSH5 defects and agents that affect the gene. The products can  
 CC also be used for gene therapy. The present sequence represents a murine  
 CC MSH5 (mMSH5) protein.  
 SQ Sequence 833 AA;

Query Match 86.0%; Score 43; DB 1; Length 833;  
 Best Local Similarity 85.7%; Pred. No. 3.43e+01;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 557 IRNGRHP 563

Qy 99 IRNGTHP 105

```

RESULT 2
ID W21737 standard; Protein; 1957 AA.
AC W21737;
DT 09-OCT-1997 (first entry)
DE Wild type rat DRG (SNS-B).
KW Rat; sensory neuron sodium channel protein; insensitive; tetrodotoxin;
KW modulator; impulse; sensory neuron; acute pain; chronic pain;
KW neuropathic pain; glia; muscle; parasympathetic nervous system;
KW enteric nervous system; central nervous system; dorsal root ganglia;
KW cranial ganglia.
OS Rattus rattus.
PN WO9701577-A1.
PD 16-JAN-1997.
PF 25-JUN-1996; G01523.
PR 28-JUN-1995; GB-013180.
PA (UNLO ) UNIV COLLEGE LONDON.
PI Akopian AN, Wood JN;
DR WPI; 97-100165/09.
DR N-PSDB; T77803.
DR New isolated mammalian sensory neuron sodium channel protein - used
PT to identify modulators of the sodium channel, partic. for the
PT treatment of pain
PS Claim 5; Page 50-58; 128pp; English.
CC The sequences given in W21737-40 represent the wild type and three
CC variant forms of a rat sensory neuron sodium channel protein which
CC is insensitive to tetrodotoxin. The proteins can be used for
CC identifying modulators of the sodium channel. Blockers of the
CC sodium channel will block or prevent the transmission of impulses
CC along sensory neurons and thereby be useful in the treatment of acute,
CC chronic or neuropathic pain. The novel protein is found only in sensory
CC neurons and not in glia, muscle or the neurons of the (para)sympathetic,
CC enteric or central nervous system. The protein is found preferably in
CC the neurons of the dorsal root ganglia or cranial ganglia.
SQ Sequence 1957 AA;

Query Match 82.0%; Score 41; DB 1; Length 1957;
Best Local Similarity 85.7%; Pred. No. 6.65e+01;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 277 IRNGTDP 283
QY 99 IRNGTHP 105

RESULT 3
ID W21740 standard; Protein; 1957 AA.
AC W21740;
DT 09-OCT-1997 (first entry)
DE Variant rat DRG (SNS-B) #3.
KW Rat; sensory neuron sodium channel protein; insensitive; tetrodotoxin;
KW modulator; impulse; sensory neuron; acute pain; chronic pain;
KW neuropathic pain; glia; muscle; parasympathetic nervous system;
KW enteric nervous system; central nervous system; dorsal root ganglia;
KW cranial ganglia.
OS Rattus rattus.
PN Key
PH Key Location/Qualifiers
FT misc_difference 297 /label= Val297Ile
FT misc_difference 298 /label= Ser298Phe
FT misc_difference 588 /label= His588Asp
FT misc_difference 757 /label= Thr757Ser
FT misc_difference 928 /label= Ser928Ile
FT misc_difference 939 /label= His939Arg
FT misc_difference 1118 /label= Arg1118Thr
FT misc_difference 1897 /label= Ile1897Ser
PN WO9701577-A1.

```

```

PD 16-JAN-1997.
PF 25-JUN-1996; G01523.
PR 28-JUN-1995; GB-013180.
PA (UNLO ) UNIV COLLEGE LONDON.
PI Akopian AN, Wood JN;
DR WPI; 97-100165/09.
DR N-PSDB; T77806.
DR New isolated mammalian sensory neuron sodium channel protein - used
PT to identify modulators of the sodium channel, partic. for the
PT treatment of pain
PS Claim 5; Page 85-93; 128pp; English.
CC The sequences given in W21737-40 represent the wild type and three
CC variant forms of a rat sensory neuron sodium channel protein which
CC is insensitive to tetrodotoxin. The proteins can be used for
CC identifying modulators of the sodium channel. Blockers of the
CC sodium channel will block or prevent the transmission of impulses
CC along sensory neurons and thereby be useful in the treatment of acute,
CC chronic or neuropathic pain. The novel protein is found only in sensory
CC neurons and not in glia, muscle or the neurons of the (para)sympathetic,
CC enteric or central nervous system. The protein is found preferably in
CC the neurons of the dorsal root ganglia or cranial ganglia. This
CC sequence contains 9 amino acid differences to the wildtype rat
CC DRG(SNS-B) (see also W21737) caused by 12 nucleotide changes.
SQ Sequence 1957 AA;

Query Match 82.0%; Score 41; DB 1; Length 1957;
Best Local Similarity 85.7%; Pred. No. 6.65e+01;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 277 IRNGTDP 283
QY 99 IRNGTHP 105

RESULT 4
ID W21739 standard; Protein; 2132 AA.
AC W21739;
DT 09-OCT-1997 (first entry)
DE Variant rat DRG (SNS-B) #2.
KW Rat; sensory neuron sodium channel protein; insensitive; tetrodotoxin;
KW modulator; impulse; sensory neuron; acute pain; chronic pain;
KW neuropathic pain; glia; muscle; parasympathetic nervous system;
KW enteric nervous system; central nervous system; dorsal root ganglia;
KW cranial ganglia.
OS Rattus rattus.
PN Key
PH Key Location/Qualifiers
FT region 586..760 /note= "Repeat unit"
FT region 586..935 /note= "Repeat region"
FT WO9701577-A1.
PN 16-JAN-1997.
PD 25-JUN-1996; G01523.
PR 28-JUN-1995; GB-013180.
PA (UNLO ) UNIV COLLEGE LONDON.
PI Akopian AN, Wood JN;
DR WPI; 97-100165/09.
DR N-PSDB; T77805.
DR New isolated mammalian sensory neuron sodium channel protein - used
PT to identify modulators of the sodium channel, partic. for the
PT treatment of pain
PS Claim 5; Page 69-78; 128pp; English.
CC The sequences given in W21737-40 represent the wild type and three
CC variant forms of a rat sensory neuron sodium channel protein which
CC is insensitive to tetrodotoxin. The proteins can be used for
CC identifying modulators of the sodium channel. Blockers of the
CC sodium channel will block or prevent the transmission of impulses
CC along sensory neurons and thereby be useful in the treatment of acute,
CC chronic or neuropathic pain. The novel protein is found only in sensory
CC neurons and not in glia, muscle or the neurons of the (para)sympathetic,
CC enteric or central nervous system. The protein is found preferably in
CC the neurons of the dorsal root ganglia or cranial ganglia. This
CC sequence represents a 2132 amino acid protein that contains a 176 amino

```

CC acid repeat inserted after amino acid 585 of wildtype rat DRG(SNS-B)  
CC (see also W21737).  
SQ Sequence 2132 AA;  
  
Query Match 82.0%; Score 41; DB 1; Length 2132;  
Best Local Similarity 85.7%; Pred. No. 6.65e+01;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Db 277 IRNGTDP 283  
| | | | |  
QY 99 IRNGTHP 105  
  
RESULT 5  
ID W08137 standard; Protein; 258 AA.  
AC W08137;  
DE 11-MAR-1997 (first entry)  
DE Human cytokine response protein CR5.  
KW Cytokine response protein; CR5; interleukin-2; IL-2;  
KW ligand-stimulated gene expression; diagnosis; therapy.  
OS Homo sapiens.  
PN W09639427-Al.  
PD 12-DEC-1996.  
PF 05-JUN-1996; U09194.  
PR 05-JUN-1995; US-461379.  
PR 05-JUN-1995; US-465585.  
PR 05-JUN-1995; US-462337.  
PR 05-JUN-1995; US-463081.  
PR 05-JUN-1995; US-462390.  
PR 05-JUN-1995; US-463074.  
PR (DART-) DARTMOUTH COLLEGE.  
PA Beadling C. Smith KA;  
PI WPI: 97-043062/04.  
DR N-PSDB: T43380.  
PT Cytokine response proteins and genes - used in the detection and  
PT therapy of diseases caused by a mutation in the CR coding region  
PS Claim 5; Page 25-26; 81pp; English.  
CC Cytokine response proteins CR1-CR8 (W08133-40) are encoded by  
CC genes (see also T43376-83) isolated from a thiol-selected  
CC interleukin-2-induced human T-cell blast cDNA library. 6 Genes  
CC (CR1, 2, 3, 5, 6, 8) are novel. CR5 expression is induced by  
CC proliferation-promoting cytokines. CR5 appears to function as a  
CC ligand-stimulated factor that facilitates mRNA expression by  
CC promoting the full elongation of mRNA transcripts. Novel agents  
CC that modify CR5 function may provide new ways of altering ligand-  
CC stimulated gene expression and thereby alter cellular function.  
CC Recombinant CR5 polypeptides can be produced in host cells.  
SQ Sequence 258 AA;  
  
Query Match 80.0%; Score 40; DB 1; Length 258;  
Best Local Similarity 57.1%; Pred. No. 9.22e+01;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
Db 106 VRDSTHP 112  
| | | | |  
QY 99 IRNGTHP 105  
  
RESULT 6  
ID W38319 standard; Protein; 258 AA.  
AC W38319;  
DE 07-JUL-1998 (first entry)  
DE Human cytokine-inducible SH2-containing (CIS) protein.  
KW Human cytokine-inducible SH2-containing; CIS; erythropoietin; EPO;  
KW tyrosine-phosphorylated interleukin; (IL)-3; signal transduction;  
KW inhibition; anaemia; ss.  
OS Homo sapiens.  
PN W09744347-Al.  
PD 27-NOV-1997.. U07477.  
PF 21-MAY-1996; WO-007477.  
PR 21-MAY-1996; WO-007477.  
PA (HARD ) HARVARD COLLEGE.  
PA (SMK ) SMITHKLINE BEECHAM CORP.

PI Dunnington DD, Frantz JD, Shoelson SE;  
DR WPI: 98-018425/02.  
DR N-PSDB: T96002.  
PT Human cytokine-inducible SH2-containing protein and related DNA -  
PT useful for diagnosis of modulators for treatment of, e.g. anaemia  
PS Claim 4; Pages 35-36; 52pp; English.  
CC This is the amino acid sequence for the human cytokine-inducible  
CC SH2-containing (CIS) protein which binds tyrosine-phosphorylated  
CC interleukin (IL)-3 or erythropoietin (EPO) receptors. When CIS is  
CC over expressed signal transduction through the receptors is inhibited.  
CC Inactivation of CIS may enhance signalling through (IL)-3 and EPO  
CC receptors. Specific inhibitors of CIS may be useful in the treatment  
CC of anaemia. Labelled CIS or its functional derivatives can be used  
CC in binding assays to determine modulators of CIS activity. Conditions  
CC associated with CIS protein deficiency can be diagnosed by assaying  
CC for the presence of the CIS gene. It can be used to treat conditions  
CC related to insufficient CIS protein function.  
SQ Sequence 258 AA;  
  
Query Match 80.0%; Score 40; DB 1; Length 258;  
Best Local Similarity 57.1%; Pred. No. 9.22e+01;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
Db 106 VRDSTHP 112  
| | | | |  
QY 99 IRNGTHP 105  
  
RESULT 7  
ID R55197 standard; Protein; 415 AA.  
AC R55197;  
DE 31-JAN-1995 (first entry)  
DE Rabbit zona pellucida ZPC protein.  
KW Rabbit; lapine; zona pellucida; ZPC; immunocontraception.  
OS Oryctolagus cuniculus.  
EH Key Location/Qualifiers  
FI protein 1 415  
FI /label= rabbit\_ZPC  
PI W09411019-A.  
PD 26-MAY-1994. U10851.  
PD 06-NOV-1993; US-973341.  
PR 09-NOV-1992; US-973341.  
PR 29-JAN-1993; US-012990.  
PA (ZONA-) ZONAGEN INC.  
PI Harris JD, Hsu KT, Podolski JS;  
DR WPI: 94-183156/22.  
DR N-PSDB: Q65607.  
PT Use of zona pellucida proteins and antibodies - for inducing  
PT reproducible transient infertility or permanent sterility in  
PT female mammals  
PS Claim 40; Page 83-84; 154pp; English.  
CC A cDNA library was prepared from mRNA isolated from ovaries removed  
CC from 5 week old rabbits. The lambda gt10 library was screened with  
CC a porcine ZPC cDNA probe (Q65606). Two positive clones were sequenced  
CC and were found to be identical except that one contained 4 extra  
CC nucleotides at the 5'-end. The determined sequence was 75%  
CC homologous to DNA coding for porcine ZPC. R55197 is the deduced amino  
CC acid sequence of rabbit ZPC.  
SQ Sequence 415 AA;  
  
Query Match 80.0%; Score 40; DB 1; Length 415;  
Best Local Similarity 57.1%; Pred. No. 9.22e+01;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
Db 196 VRTGSHP 202  
| | | | |  
QY 99 IRNGTHP 105  
  
RESULT 8  
ID W81807 standard; Protein; 415 AA.  
AC W81807;  
DE 29-JAN-1999 (first entry)

DE Rabbit ZPC protein.  
 KW ZPC; zona pellucida; infertility; sterility; immunocontraceptive;  
 KW vaccine; rabbit.  
 OS Oryctolagus cuniculus.  
 PN US5837497-A.  
 PD 17-NOV-1998.  
 PF 07-JUN-1995; 484993.  
 PR 09-NOV-1993; US-149223.  
 PR 09-NOV-1992; US-973341.  
 PR 29-JAN-1993; US-012990.  
 PR 07-JUN-1995; US-484993.  
 PA (ZONA-) ZONAGEN INC.  
 PI Harris JD;  
 DR WPI: 99-023447/02.  
 DR N-PSDB; V64789.  
 PT Isolated zona pellucida DNA from different mammals - used to develop  
 PT products which can be used for vaccination to induce transient  
 PT infertility or permanent sterility in female mammals.  
 PS Claim 5; Column 61-64; 84pp; English.  
 CC This sequence represents a rabbit ZPC protein isolated from zona  
 CC pellucida. This protein can be used in a method for specifically  
 CC inducing transient infertility or permanent sterility in a host  
 CC animal by selective vaccination with specific zona pellucida proteins  
 CC or immunocontraceptively active fragments.  
 SQ Sequence 415 AA;

Query Match 80.0%; Score 40; DB 1; Length 415;  
 Best Local Similarity 57.1%; Pred. No. 9.22e+01;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 196 VRTGSH 202  
 QY 99 IRNGTHP 105

## RESULT 9

ID R41945 standard; Protein; 135 AA.  
 AC R41945;  
 DT 10-MAY-1994 (first entry)  
 DE Serotonin receptor 5-HT2f C-terminal region.  
 KW 5-hydroxytryptamine; receptor; serotonergic; vasoconstriction;  
 KW plasmid PHD5HT2f; probe.  
 PN EP-565370-A.  
 PD 13-OCT-1993.  
 PF 07-APR-1993; 302759.  
 PR 09-APR-1992; US-864005.  
 PA (ELIL) LILLY & CO ELI.  
 PI Baez M, Kursar JD;  
 DR WPI: 93-322574/41.  
 DR N-PSDB; Q49783.  
 PT 5-Hydroxy-tryptamine receptor - used to identify drugs with  
 PT receptor activity  
 PS Disclosure: Page 17-18; 20pp; English.  
 CC The plasmid PHD5HT2f comprises the sequence Q49781 which codes for a  
 CC novel serotonin receptor. Fragments of the full-length coding  
 CC region are disclosed as suitable for use as probes to find  
 CC homologous (receptor) sequences. The preferred fragments are those  
 CC coding for the G-loop, the N-terminal and the C-terminal of the  
 CC 5-HT2f receptor (R41943-R41945, respectively).  
 SQ Sequence 135 AA;

Query Match 76.0%; Score 38; DB 1; Length 135;  
 Best Local Similarity 71.4%; Pred. No. 1.76e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 88 IRNGINP 94  
 QY 99 IRNGTHP 105

## RESULT 10

ID R41942 standard; Protein; 479 AA.  
 AC R41942;

DT 10-MAY-1994 (first entry)  
 DE Serotonin receptor 5-HT2f.  
 KW 5-hydroxytryptamine; receptor; serotonergic; vasoconstriction;  
 KW plasmid PHD5HT2f.  
 PN EP-565370-A.  
 PD 13-OCT-1993.  
 PF 07-APR-1993; 302759.  
 PR 09-APR-1992; US-864005.  
 PA (ELIL) LILLY & CO ELI.  
 PI Baez M, Kursar JD;  
 DR WPI: 93-322574/41.  
 DR N-PSDB; Q49781.  
 PT 5-Hydroxy-tryptamine receptor - used to identify drugs with  
 PT receptor activity  
 PS Claim 1; Page 14-15; 20pp; English.  
 CC The plasmid PHD5HT2f comprises the sequence Q49781 which codes for a  
 CC novel serotonin receptor. Recombinant production of the receptor  
 CC facilitates testing of compounds to ascertain the strength of their  
 CC receptor binding activity, e.g. in drug research.  
 SQ Sequence 479 AA;

Query Match 76.0%; Score 38; DB 1; Length 479;  
 Best Local Similarity 71.4%; Pred. No. 1.76e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 432 IRNGINP 438  
 QY 99 IRNGTHP 105

## RESULT 11

ID W40814 standard; Protein; 481 AA.  
 AC W40814;  
 DT 01-APR-1998 (first entry)  
 DE Human 5-HT2B receptor.  
 KW Human; serotonin receptor; 5-HT2B receptor; 5-hydroxytryptamine; 5-HT;  
 KW antibody; therapy; anxiety; depression; hypertension; migraine; autism;  
 KW compulsive disorder; schizophrenia; neurodegenerative disorder.  
 OS Homo sapiens.  
 PN US5698444-A.  
 PD 16-DEC-1997.  
 PF 23-DEC-1993; I73436.  
 PR 23-DEC-1993; US-173436.  
 PA (ELIL) LILLY & CO ELI.  
 PI Baez M, Kursar JD;  
 DR WPI: 98-051499/05.  
 DR N-PSDB; T99250.  
 PT Nucleic acid sequence encoding human serotonin receptor protein -  
 PT designated 5-HT2B, useful in screening assays for agonists and  
 PT antagonists, or to produce antibodies  
 PS Claim 1; Column 27-30; 23pp; English.  
 CC This sequence represents the human serotonin receptor, designated 5-HT2B.  
 CC Serotonin is also referred to as 5-hydroxytryptamine (5-HT). The DNA  
 CC encoding this sequence can be used for the recombinant production of  
 CC 5-HT2B, which can be used in screening assays for 5-HT2B receptor  
 CC agonists or antagonists, or to produce antibodies for therapeutic and  
 CC diagnostic applications. The agonists and antagonists are particularly of  
 CC interest in the treatment of a wide range of disorders including anxiety,  
 CC depression, hypertension, migraine, compulsive disorders, schizophrenia,  
 CC autism, neurodegenerative disorders (including Alzheimer's disease,  
 CC Parkinson's disease and Huntington's chorea), and cancer-chemotherapy  
 CC induced vomiting.  
 SQ Sequence 481 AA;

Query Match 76.0%; Score 38; DB 1; Length 481;  
 Best Local Similarity 71.4%; Pred. No. 1.76e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 433 IRNGINP 439  
 QY 99 IRNGTHP 105



RESULT 12  
ID R71033 standard; Protein; 481 AA.  
AC R71033;  
DE 02-OCT-1995 (first entry)  
DE G protein protein coupled serotonin receptor.  
KW 5-HT<sub>2</sub>; 5-HT<sub>2b</sub>; diseases; detection; diagnosis.  
OS Homo sapiens.  
PN WO9506117-A.  
PD 02-MAR-1995.  
PF 18-AUG-1994; G01813.  
PR 20-AUG-1993; GB-017349.  
PR 14-JAN-1994; GB-000597.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
PI Carey JE, Flanagan TP;  
DR WPI: 95-106846/14.  
DR N-PSDB; Q84680.  
PT New human 5-HT<sub>2</sub> receptor and nucleic acids - used in drug  
PT screening and development as well as therapy and diagnosis  
PT involving 5-HT<sub>2</sub> receptor disease states  
PS Claim 6; Page 24; 34pp; English.  
CC The sequence is the prod. of a novel gene derived from screening  
CC to a human small intestine cDNA library. The gene belongs  
CC to the G protein coupled receptor family having strong homology to  
CC the rat and mouse 5-HT<sub>2b</sub> receptor sequences. The DNA can be used for  
CC treating diseases associated with excess serotonin receptor  
CC activation.  
CC See also R71032.  
SQ Sequence 481 AA;  
Query Match 76.0%; Score 38; DB 1; Length 481;  
Best Local Similarity 71.4%; Pred. No. 1.76e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Db 433 IRNGINP 439  
QY 99 IRNGTHP 105  
|||||:  
RESULT 13  
ID R5482 standard; Protein; 504 AA.  
AC R5482; 1994 (first entry)  
DE Mouse brain 5HT<sub>2C</sub> serotoninergic receptor.  
KW Serotonin; 5-hydroxytryptamine; serotoninergic receptor activity;  
KW neuromodulator; mouse; murine; 5HT<sub>2C</sub>.  
OS Mus musculus.  
PN FR2696749-A.  
PD 15-APR-1994.  
PF 14-OCT-1992; 012280.  
PR 14-OCT-1992; FR-012280.  
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
PI Maroteaux L;  
DR WPI: 94-146426/18.  
DR P-PSDB; R54682.  
PT New protein 5HT<sub>2C</sub> with serotoninergic receptor activity - and  
PT related nucleic acid, anti-sense sequences, diagnostic probes,  
PT recombinant cells and therapeutic receptor ligands and modulators  
PS Claim 1; Page 18-19; 26pp; French.  
CC Degenerate oligonucleotides corresp. to a conserved part of  
CC transmembrane regions 6 and 7 of known serotoninergic receptor  
CC used in PCR amplification of mouse genomic DNA. Amplification  
CC products were sequenced and one which had some homology with 5HT<sub>2</sub>  
CC receptors was synthesised in vitro in labelled form. A mouse brain  
CC cDNA library was screened with the labelled probe and the positive  
CC clone NP75 was selected. This clone (Q84657) contains an open  
CC reading frame encoding the 504 amino acid protein R54682.  
SQ Sequence 504 AA;  
Query Match 76.0%; Score 38; DB 1; Length 504;  
Best Local Similarity 71.4%; Pred. No. 1.76e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Db 432 IRNGINP 438

QY 99 IRNGTHP 105  
|||||:  
RESULT 14  
ID W81351 standard; Protein; 550 AA.  
AC W81351;  
DE 24-MAR-1999 (first entry)  
DE Human guanine nucleotide exchange factor CSB5 variant.  
KW Human; guanine nucleotide exchange factor; GEF; CSB5; heart hypertrophy;  
KW heart failure; ischaemia; arrhythmia; hypertension; atherosclerosis;  
KW restenosis; chronic inflammation; acute inflammation; cerebral stroke;  
KW rheumatoid arthritis; multiple sclerosis; bowel disease; diabetes;  
KW cancer.  
OS Homo sapiens.  
PN EP-882792-A2.  
PD 09-DEC-1998.  
PF 28-MAY-1998; 304207.  
PR 26-DEC-1997; EP-403166.  
PR 06-JUN-1997; EP-401277.  
PR 26-DEC-1997; EP-403165.  
PA (SMIK ) SMITHKLINE BEECHAM CORP.  
PI Brill ANA, Calmels TGG, Hurle MR, Leger IM, Souchet ML,  
PI Tourtelier LNP;  
DR WPI: 99-011645/02.  
DR N-PSDB; V88453.  
PT New human CSB5 polypeptide and related nucleic acid, vectors and  
PT transformed cells - and corresponding antibodies, agonists and  
PT antagonists, useful for diagnosis, treatment and prevention of  
PT cardiac hypertrophy, hypertension, cancer, multiple sclerosis,  
PT diabetes, inflammation etc.  
PS Claim 4; Page 15; 27pp; English.  
CC The present sequence is a human guanine nucleotide exchange factor  
CC (GEF) family protein variant designated CSB5var. Host cells containing  
CC vectors with CSB5 nucleotide sequences are used to produce recombinant  
CC CSB5 which is used to screen for specific (antagonists (potential  
CC therapeutic agents) and to generate AB (including induction of a  
CC protective immune response). CSB5 is involved in e.g. cytoskeletal  
CC organisation; cell growth; proliferation; migration and adhesion;  
CC membrane trafficking; vesicle transport; and apoptosis. CSB5 and its  
CC (antagonists are used for treatment or prevention of heart hypertrophy,  
CC heart failure, ischaemia, arrhythmia, hypertension, atherosclerosis,  
CC restenosis, chronic or acute inflammation, cerebral stroke, rheumatoid  
CC arthritis, multiple sclerosis, bowel disease, diabetes, and cancer.  
SQ Sequence 550 AA;  
Query Match 76.0%; Score 38; DB 1; Length 550;  
Best Local Similarity 57.1%; Pred. No. 1.76e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Db 368 VRGGTQP 374  
QY 99 IRNGTHP 105  
|||||:  
RESULT 15  
ID W81349 standard; Protein; 580 AA.  
AC W81349;  
DE 24-MAR-1999 (first entry)  
DE Human guanine nucleotide exchange factor CSB5.  
KW Human; guanine nucleotide exchange factor; GEF; CSB5; heart hypertrophy;  
KW heart failure; ischaemia; arrhythmia; hypertension; atherosclerosis;  
KW restenosis; chronic inflammation; acute inflammation; cerebral stroke;  
KW rheumatoid arthritis; multiple sclerosis; bowel disease; diabetes;  
KW cancer; ss.  
OS Homo sapiens.  
PN EP-882792-A2.  
PD 09-DEC-1998.  
PF 28-MAY-1998; 304207.  
PR 26-DEC-1997; EP-403166.  
PR 06-JUN-1997; EP-401277.  
PR 26-DEC-1997; EP-403165.

PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM LAB PHARM.  
 PI Brill AM, Calmels TGG, Hurle MR, Leger IM, Souchet ML,  
 PI Tourtellier LNP;  
 DR WPI: 99-011649/02.  
 DR N-PSDB; V68451.  
 PT New human CSB5 polypeptide and related nucleic acid, vectors and  
 PT transformed cells - and corresponding antibodies, agonists and  
 PT antagonists, useful for diagnosis, treatment and prevention of  
 PT cardiac hypertrophy, hypertension, cancer, multiple sclerosis,  
 PT diabetes, inflammation etc.  
 PS Claim 4, Page 13; 27pp; English.  
 CC The present sequence represents human guanine nucleotide exchange factor  
 CC (GEF) family protein designated CSB5. Host cells containing vectors with  
 CC CSB5 nucleotide sequences are used to produce recombinant CSB5 which is  
 CC used to screen for specific (ant)agonists (potential therapeutic agents)  
 CC and to generate Ab (including induction of a protective immune response).  
 CC CSB5 is involved in e.g. cytoskeletal organisation; cell growth,  
 CC proliferation, migration and adhesion; membrane trafficking; vesicle  
 CC transport; and apoptosis. CSB5 and its (ant)agonists are used for  
 CC treatment or prevention of heart hypertrophy, heart failure, ischaemia,  
 CC arrhythmia, hypertension, atherosclerosis, restenosis, chronic or acute  
 CC inflammation, cerebral stroke, rheumatoid arthritis, multiple sclerosis,  
 CC bowel disease, diabetes, and cancer.  
 SQ Sequence 580 AA;

Query Match 76.0%; Score 38; DB 1; Length 580;  
 Best Local Similarity 57.1%; Pred. No. 1.76e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 398 VRGGTQP 404

OY 99 IRNGTHP 105

Search completed: Wed May 10 12:59:02 2000  
 Job time : 7 secs.

\*\*\*\*\*  
 W P E R L H  
 \*\*\*\*\* (TY)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MParch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 12:59:18 2000; MasPar time 45.02 Seconds  
 Tabular output not generated.  
 2.015 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (99-105) from US09376430A.pep (14 of 25)  
 Perfect Score: 50  
 Sequence: 1 IRNGTHP 7

Scoring table: PAM 150  
 Gap 11

Searched: 131253 seqs, 12956647 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: 1:5A\_COMB 2:5B\_COMB 3:PCT\_COMB 4:backfiles1

Statistics: Mean 13.286; Variance 29.199; scale 0.455

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	40	80.0	257	2	US-08-918- Sequence 3, Applicatio	3.82e+01
2	40	80.0	258	2	US-08-462- Sequence 10, Applicati	3.82e+01
3	40	80.0	258	2	US-08-461- Sequence 10, Applicati	3.82e+01
4	40	80.0	258	2	US-08-463- Sequence 10, Applicati	3.82e+01
5	40	80.0	415	2	US-08-484- Sequence 8, Applicatio	3.82e+01
6	40	80.0	415	2	US-08-480- Sequence 8, Applicatio	3.82e+01
7	40	80.0	415	2	US-08-484- Sequence 8, Applicatio	3.82e+01
8	40	80.0	415	2	US-08-484- Sequence 8, Applicatio	3.82e+01
9	38	76.0	479	1	US-08-416- Sequence 2, Applicatio	7.45e+01
10	38	76.0	481	1	US-08-173- Sequence 2, Applicatio	7.45e+01
11	38	76.0	481	2	US-08-748- Sequence 8, Applicatio	7.45e+01
12	38	76.0	623	1	US-08-734- Sequence 2, Applicatio	7.45e+01
13	37	74.0	171	2	US-08-853- Sequence 54, Applicati	1.04e+02
14	37	74.0	806	1	US-08-451- Sequence 6, Applicatio	1.04e+02
15	37	74.0	1157	2	US-08-533- Sequence 5, Applicatio	1.04e+02
16	37	74.0	1157	2	US-08-379- Sequence 5, Applicatio	1.04e+02
17	37	74.0	1169	1	US-08-543- Sequence 2, Applicatio	1.04e+02
18	37	74.0	1169	2	US-08-880- Sequence 2, Applicatio	1.04e+02
19	37	74.0	1169	2	US-08-880- Sequence 2, Applicatio	1.04e+02
20	35	70.0	50	2	US-08-956- Sequence 6, Applicatio	1.98e+02
21	35	70.0	50	2	US-08-956- Sequence 6, Applicatio	1.98e+02
22	35	70.0	50	1	US-08-656- Sequence 6, Applicatio	1.98e+02
23	35	70.0	50	2	US-08-777- Sequence 10, Applicati	1.98e+02

24	35	70.0	50	2	US-08-777- Sequence 20, Applicati	1.98e+02
25	35	70.0	50	1	US-08-377- Sequence 21, Applicati	1.98e+02
26	35	70.0	50	1	US-08-377- Sequence 10, Applicati	1.98e+02
27	35	70.0	50	1	US-08-556- Sequence 5, Applicatio	1.98e+02
28	35	70.0	50	2	US-08-777- Sequence 21, Applicati	1.98e+02
29	35	70.0	50	1	US-08-377- Sequence 12, Applicati	1.98e+02
30	35	70.0	50	2	US-08-777- Sequence 12, Applicati	1.98e+02
31	35	70.0	261	3	PCT-US96-1 Sequence 25, Applicati	1.98e+02
32	35	70.0	277	3	PCT-US96-1 Sequence 8, Applicatio	1.98e+02
33	35	70.0	462	2	US-08-752- Sequence 7, Applicatio	1.98e+02
34	35	70.0	464	3	PCT-US96-1 Sequence 18, Applicatio	1.98e+02
35	35	70.0	465	2	US-08-752- Sequence 5, Applicatio	1.98e+02
36	35	70.0	479	3	US-08-807- Sequence 12, Applicati	1.98e+02
37	35	70.0	479	3	PCT-US96-1 Sequence 7, Applicatio	1.98e+02
38	35	70.0	496	1	US-08-865- Sequence 4, Applicatio	1.98e+02
39	35	70.0	612	2	US-08-752- Sequence 11, Applicati	1.98e+02
40	35	70.0	1184	1	US-08-446- Sequence 20, Applicati	1.98e+02
41	35	70.0	1184	2	US-08-064- Sequence 20, Applicati	1.98e+02
42	35	70.0	1187	2	US-09-003- Sequence 8, Applicatio	1.98e+02
43	35	70.0	1187	3	PCT-US95-1 Sequence 8, Applicatio	1.98e+02
44	35	70.0	1187	1	US-08-097- Sequence 13, Applicati	1.98e+02
45	35	70.0	1367	2	US-08-249- Sequence 2, Applicatio	1.98e+02

## ALIGNMENTS

RESULT 1  
 ID US-08-918-206-3 STANDARD; PRT; 257 AA.  
 AC xxxxxx  
 XX  
 DT

Sequence 3, Application US/08918206

Sequence 3, Application US/08918206  
 Patent No. 5919661

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Guegler, Karl

APPLICANT: Corley, Neil C.

APPLICANT: Shah, Purvi

TITLE OF INVENTION: CYTOKINE INDUCIBLE REGULATORY

TITLE OF INVENTION: PROTEIN

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Dr.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/918,206

FILING DATE: Filed Herewith

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0372 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 257 amino acids

```
Query Match      80.0%; Score 40; DB 2; Length 258;
Best Local Similarity 57.1%; Pred. No. 3.82e+01;
Matches      4; Conservative 3; Mismatches 0; Indels
```

XXXXXX



```
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: United States of America
CC ZIP: 60606-6402
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/480,150A
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/149,223
CC FILING DATE: 09-NOV-1993
CC APPLICATION NUMBER: 08/012,990
CC FILING DATE: 29-JAN-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/973,341
CC FILING DATE: 09-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Clough, David W.
CC REGISTRATION NUMBER: 36,107
CC REFERENCE/DOCKET NUMBER: 31745
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312/474-6653
CC TELEFAX: 312/474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 415 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 415 AA; 44987 MW; 849304 CN;

Query Match      80.0%; Score 40; DB 2; Length 415;
Best Local Similarity 57.1%; Pred.No. 3.82e+01;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db    196 VRTGSHP 202
Qy    99 IRNGTHP 105

RESULT 7
ID ID US-08-484-596A-8 STANDARD; PRT; 415 AA.
XX AC xxxxxx
XX XX
XX DT
XX DE
XX DE Sequence 8, Application US/08484596A
XX XX
XX CC Sequence 8, Application US/08484596A
XX CC Patent No. 5981228
XX CC GENERAL INFORMATION:
XX CC APPLICANT: Harris Ph.D., Jeffrey D.
XX CC APPLICANT: Hsu, Kuang T.
XX CC APPLICANT: Podolski, Joseph S.
XX CC TITLE OF INVENTION: Materials and Methods for Immunocontraception
XX CC NUMBER OF SEQUENCES: 59
XX CC CORRESPONDENCE ADDRESS:
XX CC ADDRESSSEE: Marshall, O'Toole, Gerstein, Murray & Borun
XX CC STREET: 6300 Sears Tower, 233 South Wacker Drive
XX CC CITY: Chicago
XX CC STATE: Illinois
XX CC COUNTRY: United States of America
XX CC ZIP: 60606-6402
XX CC COMPUTER READABLE FORM:
XX CC MEDIUM TYPE: Floppy disk
XX CC COMPUTER: IBM PC compatible
```

US-09-376-430-2-14.rai

Thu May 11 06:49:55 2000

CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: 07/973,341  
 CC FILING DATE: 03-NOV-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Clough, David W.  
 CC REGISTRATION NUMBER: 36,107  
 CC REFERENCE/DOCKET NUMBER: 31745  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 312/474-0448  
 CC TELEX: 25-3856  
 CC INFORMATION FOR SEQ ID NO: 8:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 415 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 415 AA; 41987 MW; 849304 CN;  
 CC  
 CC Query Match 80.0%; Score 40; DB 2; Length 415;  
 CC Best Local Similarity 57.1%; Pred. No. 3.82e+01;  
 CC Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 Db 196 VRTGSH 202  
 QY 99 IRNGTHP 105  
 CC  
 CC RESULT 9  
 CC ID US-08-416-788-2 STANDARD; PRT; 479 AA.  
 CC XX  
 CC AC xxxxxx  
 CC DT  
 CC XX  
 CC DE  
 CC SEQUENCE 2, Application US/08416788  
 CC Patent No. 5780245  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Marcoteaux, Luc  
 CC TITLE OF INVENTION: No. 5780245el Polypeptides Having a Serotonin  
 CC TITLE OF INVENTION: Receptor Activity, Nucleic Acids Coding for These  
 CC TITLE OF INVENTION: Polypeptides and Uses  
 CC NUMBER OF SEQUENCES: 9  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Rhone-Poulenc Rorsr Inc.  
 CC STREET: 500 Arcoia Road, 3C43  
 CC CITY: Collegeville  
 CC STATE: PA  
 CC COUNTRY: USA  
 CC ZIP: 19426-0107  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent in Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/416,788  
 CC FILING DATE:  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: WO PCT/FR93/01012  
 CC FILING DATE: 13-OCT-1993  
 CC PRIOR APPLICATION DATA: FR 92-12280  
 CC FILING DATE: 14-OCT-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Smith, Julie K.  
 CC REGISTRATION NUMBER: 38,619  
 CC REFERENCE/DOCKET NUMBER: EX92008-US  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (610)454-3839

CC TELEFAX: (610)454-3808  
 CC INFORMATION FOR SEQ ID NO: 2:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 479 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 479 AA; 53654 MW; 1256982 CN;  
 CC  
 CC Query Match 76.0%; Score 38; DB 1; Length 479;  
 CC Best Local Similarity 71.4%; Pred. No. 7.45e+01;  
 CC Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 Db 432 IRNGINP 438  
 QY 99 IRNGTHP 105  
 CC  
 CC RESULT 10  
 CC ID US-08-173-436A-2 STANDARD; PRT; 481 AA.  
 CC XX  
 CC AC xxxxxx  
 CC DT  
 CC XX  
 CC DE  
 CC SEQUENCE 2, Application US/08173436A  
 CC Patent No. 5698444  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Baez, Melvyn  
 CC APPLICANT: Karsak, Jonathan D.  
 CC TITLE OF INVENTION: SEROTONIN RECEPTOR PROTEIN AND RELATED  
 CC TITLE OF INVENTION: NUCLEIC ACID COMPOUNDS  
 CC NUMBER OF SEQUENCES: 4  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Eli Lilly and Company  
 CC STREET: Lilly Corporate Center/Patent Division  
 CC CITY: Indianapolis  
 CC STATE: IN  
 CC COUNTRY: US  
 CC ZIP: 46285  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patent in Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/173,436A  
 CC FILING DATE: 23-DEC-1993  
 CC CLASSIFICATION: 435  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Gaylo, Paul J.  
 CC REGISTRATION NUMBER: 36,808  
 CC REFERENCE/DOCKET NUMBER: X-9367  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 317-276-0756  
 CC TELEFAX: 317-276-3861  
 CC INFORMATION FOR SEQ ID NO: 2:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 481 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 481 AA; 54297 MW; 1282850 CN;  
 CC  
 CC Query Match 76.0%; Score 38; DB 1; Length 481;  
 CC Best Local Similarity 71.4%; Pred. No. 7.45e+01;  
 CC Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 Db 433 IRNGINP 439  
 QY 99 IRNGTHP 105





US-08-376-430-2-14.ra1

Thu May 11 06:49:55 2000

CC APPLICANT: Wong, K.K.; Saffer, J.D.  
 CC TITLE OF INVENTION: Of A Salmonella Sequence, Methods Of Detection  
 CC TITLE OF INVENTION: Of A Salmonella Sequence, Methods Of Detection  
 CC TITLE OF INVENTION: Of A Salmonella Sequence, Methods Of Detection  
 CC NUMBER OF SEQUENCES: 67  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Paul W. Zimmerman  
 CC ADDRESSEE: Intellectual Property Services  
 CC ADDRESSEE: Battelle Memorial Institute  
 CC ADDRESSEE: PNL P.O. Box 999  
 CC STREET: Washington Way  
 CC CITY: Richland  
 CC STATE: Washington  
 CC COUNTRY: U.S.A.  
 CC ZIP: 99352  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage  
 CC COMPUTER: IBM PC/XT/AT  
 CC OPERATING SYSTEM: MS-DOS  
 CC SOFTWARE: Word Processor (Wordperfect 5.1)  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/853.659A  
 CC FILING DATE: Unknown  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: none  
 CC FILING DATE: n/a  
 CC INFORMATION FOR SEQ ID NO: 54:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 171 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC SEQUENCE 171 AA; 19340 MW; 140525 CN;  
 CC  
 CC Query Match 74.0%; Score 37; DB 2; Length 171;  
 CC Best Local Similarity 57.1%; Pred. No. 1.04e+02;  
 CC Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 Db 76 IRHGARP 82  
 QY 99 IRNGTHP 105  
 RESULT 14  
 ID US-08-451-715A-6 STANDARD; PRT; 806 AA.  
 AC xxxxxx  
 CC  
 CC Sequence 6, Application US/08451715A  
 CC Patent No. 5801013  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Tao, Jianshi  
 CC APPLICANT: Qui, Yan  
 CC APPLICANT: Houman, Fariba  
 CC APPLICANT: Shen, Xiaoyu  
 CC APPLICANT: Schimmel, Paul R.  
 CC TITLE OF INVENTION: Helicobacter Aminoacyl-tRNA Synthetase  
 CC TITLE OF INVENTION: Proteins, Nucleic Acids and Strains Comprising  
 CC NUMBER OF SEQUENCES: 67  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 CC STREET: Two Militia Drive  
 CC CITY: Lexington  
 CC STATE: Massachusetts  
 CC COUNTRY: USA  
 CC ZIP: 02173  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/451,715A  
 CC FILING DATE: 26-MAY-1995  
 CC CLASSIFICATION: 435  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Brook, David E.  
 CC REGISTRATION NUMBER: 22,592  
 CC REFERENCE/DOCKET NUMBER: CPI94-25  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 617-861-6240  
 CC TELEFAX: 617-861-9540  
 CC INFORMATION FOR SEQ ID NO: 6:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 806 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 806 AA; 93042 MW; 3270659 CN;  
 CC  
 CC Query Match 74.0%; Score 37; DB 1; Length 806;  
 CC Best Local Similarity 57.1%; Pred. No. 1.04e+02;  
 CC Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 Db 87 IRKGHP 93  
 QY 99 IRNGTHP 105  
 RESULT 15  
 ID US-08-532-547-5 STANDARD; PRT; 1157 AA.  
 AC xxxxxx  
 CC  
 CC Sequence 5, Application US/08532547  
 CC Patent No. 5861543  
 CC GENERAL INFORMATION:  
 CC APPLICANT: LAMBERT, BART  
 CC APPLICANT: JANSSENS, STEFAN  
 CC APPLICANT: VAN AUDENHOVE, KATRIEN  
 CC APPLICANT: PEEROEN, MARNIX  
 CC APPLICANT: VAN RIE, JEROEN  
 CC APPLICANT: VAN AARSSEN, ROEL  
 CC TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND THEIR  
 CC TITLE OF INVENTION: INSECTICIDAL PROTEINS.  
 CC NUMBER OF SEQUENCES: 9  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP  
 CC STREET: P.O. Box 747  
 CC CITY: Falls Church  
 CC STATE: Virginia  
 CC COUNTRY: USA  
 CC ZIP: 22040-0747  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/532,547  
 CC FILING DATE: 06-DEC-1996  
 CC CLASSIFICATION: 800  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: SVENSSON, LEONARD R.  
 CC REGISTRATION NUMBER: 30,330  
 CC REFERENCE/DOCKET NUMBER: 2121-109P  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (703) 205-8000

CC TELEFAX: (703) 205-8050  
 CC TELEX: 248345  
 CC INFORMATION FOR SEQ ID NO: 5:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 1157 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 SQ SEQUENCE 1157 AA; 12975 MW; 6884754 CN;

Query Match 74.0%; Score 37; DB 2; Length 1157;  
 Best Local Similarity 66.7%; Pred. No. 1.04e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1091 IRDGAH 1096  
 QY 99 IRNGTH 104

Search completed: Wed May 10 13:00:12 2000  
 Job time : 54 secs.

Release 3.1a John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

```

MParch_pp      protein - protein database search, using Smith-Waterman algorithm
Run on:        Wed May 10 12:58:30 2000;   MasPar time 3.86 Seconds
              85.493 Million cell updates/sec
Tabular output not generated.

```

```

Title:
Description:
Perfect Score:
Sequence:
1 IRNGTHP 7
>US-09-376-430-2
(99-105) from US09376430A.pep (14 of 25)
50

```

Scoring table: PAM 150  
Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

```
Database:
pir62
1:pir1 2:pir2 3:pir3 4:pir4
```

Statistics: Mean 19.327; Variance 18.139; scale 1.065

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	45	90.0	819	C71544	leucine--tRNA ligase	4.22e-01
2	45	90.0	820	C72113	leucine--tRNA ligase	4.22e-01
3	45	90.0	832	S20752	DNA-directed DNA poly	4.22e-01
4	43	86.0	97	S08602	hypothetical protein	1.50e+00
5	43	86.0	269	A61182	hypothetical protein	1.50e+00
6	41	82.0	937	B69068	leucine--tRNA ligase	5.14e+00
7	41	82.0	1021	T05108	hypothetical protein	5.14e+00
8	41	82.0	1957	S68453	sodium channel protei	5.14e+00
9	40	80.0	79	JQ0312	hypothetical 8.9k pro	9.36e+00
10	40	80.0	257	S55551	cytokine-inducible pr	9.36e+00
11	40	80.0	415	S70401	zona pellucida glycop	9.36e+00
12	40	80.0	519	S39893	rnc Protein - Rhodop	9.36e+00
13	40	80.0	583	T04531	nine-tis-epoxycaroten	9.36e+00
14	40	80.0	832	S71795	DNA-directed DNA poly	9.36e+00
15	40	80.0	832	JDVLYB	DNA-directed DNA poly	9.36e+00
16	40	80.0	832	JDVYAL	DNA-directed DNA poly	9.36e+00
17	40	80.0	832	S47406	DNA-directed DNA poly	9.36e+00
18	39	78.0	365	T00319	hypothetical 43.5K pr	1.69e+01
19	39	78.0	393	JVBEG6	protein kinase (EC 2	1.69e+01
20	38	76.0	427	S30358	ribose-phosphate pyro	1.69e+01
21	38	76.0	84	S03384	hypothetical protein	3.00e+01
22	38	76.0	166	T139446	hypothetical protein	3.00e+01
23	38	76.0	211	A45928	colanic acid capsular	3.00e+01

24	38	76.0	287	1	QQV22	mRNA capping enzyme s	3.00e-01
25	38	76.0	287	2	I36847	N2L protein - variola	3.00e-01
26	38	76.0	287	3	I42516	D12L protein - vaccin	3.00e-01
27	38	76.0	413	1	ALWT3	alpha-amylase (EC 3.2	3.00e-01
28	38	76.0	416	2	T00137	portal protein - Stap	3.00e-01
29	38	76.0	437	2	I14956	alpha-amylase (EC 3.2	3.00e-01
30	38	76.0	439	2	T02956	alpha-amylase (EC 3.2	3.00e-01
31	38	76.0	479	2	S23562	serotonin receptor 2B	3.00e-01
32	38	76.0	481	2	S43687	serotonin receptor 2B	3.00e-01
33	38	76.0	481	2	S49422	serotonin receptor 2B	3.00e-01
34	38	76.0	504	2	S27269	serotonin receptor 2-	3.00e-01
35	38	76.0	595	2	T04438	hypothetical protein	3.00e-01
36	38	76.0	605	2	T07133	nine cis-epoxycaroten	3.00e-01
37	38	76.0	684	2	T13401	NADH dehydrogenase -	3.00e-01
38	38	76.0	818	2	S62790	mismatch DNA recogni	3.00e-01
39	38	76.0	835	2	A53286	DNA mismatch repair p	3.00e-01
40	38	76.0	861	2	G64087	DNA mismatch repair p	3.00e-01
41	38	76.0	867	1	RVRECV	RNA-directed RNA poly	3.00e-01
42	38	76.0	929	2	A32495	rep-1 protein, form A	3.00e-01
43	38	76.0	1126	2	JC4019	DNA mismatch repair p	3.00e-01
44	38	76.0	1137	2	A33507	hypothetical protein	3.00e-01
45	38	76.0	2717	2	A34203	DNA-binding protein P	3.00e-01

## ALIGNMENTS

```

RESULT      1
ENTRY
  TITLE      C71544 #type complete
  ORGANISM   leucine--trna ligase (EC 6.1.1.4) - Chlamydia trachomatis
             (serotype D, strain UW3/Cx)
             #formal_name Chlamydia trachomatis
  DATE      13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change
             16-Jul-1999
  ACCESSIONS C71544
  REFERENCE  A71570
  #authors   Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe,
             R.; Aravind, L.; Mitchell, W.P.; Olinger, L.; Tatusov,
             R.L.; Zhao, Q.; Koonin, E.V.; Davis, R.W.
  #journal   Science (1998) 282:754-759
  #title     Genome sequence of an obligate intracellular pathogen of
             humans, Chlamydia trachomatis.
  #cross-references WUID:99000809
  #accession C71544
             #status      Preliminary
             #molecule_type DNA
             ##residues    1-819 #label ARN
             ##cross-references GB:AE001294; GB:AE001273; NID:g3328609;
             PIDN:AA667801.1; PID:g3328616
             ##experimental_source serotype D, strain UW-3/Cx

GENETICS
#gene       leuS
#superfamily leucine--trna ligase
#KEYWORDS   aminocacyl--trna synthetase; ligase; protein biosynthesis
SUMMARY
#length 819 #molecular_weight 92902 #checksum 8525

Query Match          90.0%; Score 45; DB 2; Length 819;
Best Local Similarity 85.7%; Pred. No. 4.22e-01;
Matches              6; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Db      90  IRTGTHP 96
      || ||||
QY      99  IRTGTHP 105

RESULT      2
ENTRY
  TITLE      C72113 #type complete
  ORGANISM   leucine--trna ligase (EC 6.1.1.4) - Chlamydia pneumoniae
             (strain CWL029)
             #formal_name Chlamydia pneumoniae
  DATE      23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change
             16-Jul-1999
  ACCESSIONS C72113
  REFERENCE  A72000

```

```

#authors      Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.;
#journal      Olinger, L.; Grimwood, J.; Davis, R.W.; Stephens, R.S.
#title        Comparative genomes of Chlamydia pneumoniae and C.
#cross-references EMBL:99206606
#accession     C72113
#status        preliminary
#molecule_type DNA
#residues      1-820 #label ARN
#cross-references GB:AE001602; GB:AE001363; NID:94376416;
#experimental_source strain CWL029
GENETICS
#gene          leus
CLASSIFICATION #superfamily leucine--trna ligase
KEYWORDS       aminocyl--trna synthetase; ligase; protein biosynthesis
SUMMARY        #length 820 #molecular-weight 93965 #checksum 1601
Query Match    90.0%; Score 45; DB 2; Length 820;
Best Local Similarity 85.7%; Pred. No. 4.22e-01;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 90 IRTGTHP 96
QY 99 IRNGTHP 105

RESULT 3
ENTRY   S20752 #type complete
TITLE   DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus
        (subtype ayw, patient CI)
ORGANISM #formal_name hepatitis B virus, HBV
#variety subtype ayw, patient CI
DATE     20-Feb-1995 #sequence_revision 22-Nov-1996 #text_change
        11-Jun-1999
ACCESSIONS S20752
REFERENCE   S20745
#authors    Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Mellis, A.;
            Porru, A.
#submission submitted to the EMBL Data Library, March 1992
#description Sequence analysis of HBV genomes isolated from patients with
            HBsAg negative chronic liver disease.
#accession  S20752
#molecule_type DNA
#residues   1-832 #label LAI
#cross-references EMBL:X65258; NID:959434; PIDN:CAA46356.1; PID:959438
#experimental_source subtype ayw, patient CI
GENETICS
#gene          P
CLASSIFICATION #superfamily hepatitis virus DNA-directed DNA polymerase
KEYWORDS       DNA biosynthesis; nucleotidyltransferase
SUMMARY        #length 832 #molecular-weight 93871 #checksum 1001
Query Match    90.0%; Score 45; DB 1; Length 832;
Best Local Similarity 85.7%; Pred. No. 4.22e-01;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 232 IRAGTHP 238
QY 99 IRNGTHP 105

RESULT 4
ENTRY   S08602 #type complete
TITLE   hypothetical protein 1 - phage T4
ORGANISM #formal_name phage T4
DATE     02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change
        09-Sep-1997
ACCESSIONS S08602
REFERENCE   S07190
#authors    Stoeberg, B.M.; Hahne, S.; Mathews, C.Z.; Mathews, C.K.;
            Rand, K.N.; Gait, M.J.

```

```

#journal      EMBO J. (1986) 5:2031-2036
#title        The bacteriophage T4 gene for the small subunit of
            ribonucleotide reductase contains an intron.
#cross-references MUID:87004574
#accession    S08602
#status        preliminary
#molecule_type DNA
#residues      1-97 #label SUO
#cross-references EMBL:X04140; NID:915341; PID:915343
SUMMARY        #length 97 #molecular-weight 11330 #checksum 5796
Query Match    86.0%; Score 43; DB 2; Length 97;
Best Local Similarity 71.4%; Pred. No. 1.50e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 26 VKNGTHP 32
QY 99 IRNGTHP 105

RESULT 5
ENTRY   A61182 #type complete
TITLE   hypothetical protein (nrdb intron) - phage RB3
ORGANISM #formal_name phage RB3
DATE     03-May-1994 #sequence_revision 03-May-1994 #text_change
        09-Sep-1997
ACCESSIONS A61182; S29927
REFERENCE   A61182
#authors    Eddy, S.R.; Gold, L.
#journal     Genes Dev. (1991) 5:1032-1041
#title       The phage T4 hrdb intron: a deletion mutant of a version
            found in the wild.
#cross-references MUID:91257570
#accession  A61182
#molecule_type DNA
#residues   1-269 #label EDD
#cross-references GB:X59078; NID:915180; PID:9579158
GENETICS
#start_codon TTG
SUMMARY      #length 269 #molecular-weight 30814 #checksum 9434
Query Match    86.0%; Score 43; DB 2; Length 269;
Best Local Similarity 71.4%; Pred. No. 1.50e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 198 VKNGTHP 204
QY 99 IRNGTHP 105

RESULT 6
ENTRY   B69068 #type complete
TITLE   leucine--trna ligase (EC 6.1.1.4) - Methanobacterium
            thermoautotrophicum (strain Delta H)
ALTERNATE_NAMES leucyl--trna synthetase
ORGANISM         #formal_name Methanobacterium thermoautotrophicum
DATE             05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
        16-Jul-1999
ACCESSIONS B69068
REFERENCE   A69000
#authors    Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.;
            Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.;
            Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.;
            Lumm, W.; Pothier, B.; Qiu, D.; Spadafora, R.; Vicaire, R.;
            Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiawani, N.; Caruso,
            A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.;
            McDougall, S.; Shimer, G.; Goyal, A.; Pietrokowski, S.;
            Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling,
            J.; Reeve, J.N.
#journal     J. Bacteriol. (1997) 179:7135-7155
#title       Complete genome sequence of Methanobacterium
            thermoautotrophicum Delta H: functional analysis and
            comparative genomics.

```

```

#cross-references MUID:98037514
#accession B69068
#status preliminary; nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-937 #label MTH
#cross-references GB:AE000911; GB:AE000666; NID:g2622623;
#cross-references PIDN:AB85983.1; PID:g2622626
#experimental_source strain Delta H
GENETICS
#gene MTH1508
#start_codon GTG
CLASSIFICATION
#superfamily valine--trna ligase
KEYWORDS aminoacyl--trna synthetase; ligase; protein biosynthesis
#length 937 #molecular-weight 108030 #checksum 9524
SUMMARY
Query Match 82.0%; Score 41; DB 2; Length 937;
Best Local Similarity 57.1%; Pred. No. 5.14e+00;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 165 VRKGAP 171
QY 99 IRNGTHP 105
RESULT 7
ENTRY #type complete
TITLE hypothetical protein F28M20.190 - Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
cross
DATE 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change
23-Jul-1999
ACCESSION T05108
REFERENCE T05108
#authors Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Hobeisel, J.; Mewes, H.W.; Mayer, K.F.X.; Schueller, C.
#submission submitted to the Protein Sequence Database, November 1998
#accession T05108
#molecule_type DNA
#residues 1-1021 #label BEV
#cross-references EMBL:AL031004
#experimental_source cultivar Columbia; BAC clone F28M20
GENETICS
#map_position 4
#introns 27/3; 169/3; 310/3; 395/3; 500/1; 697/3; 820/3; 918/3
#note F28M20.190
#length 1021 #molecular-weight 115721 #checksum 5344
SUMMARY
Query Match 82.0%; Score 41; DB 2; Length 1021;
Best Local Similarity 71.4%; Pred. No. 5.14e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 237 VRNGTRP 243
QY 99 IRNGTHP 105
RESULT 8
ENTRY #type complete
TITLE sodium channel protein SNS - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change
20-Aug-1999
ACCESSION S68453
REFERENCE S68453
#authors Akopian, A.N.; Sivillotti, L.; Wood, J.N.
#journal Nature (1996) 379:257-262
#title A tetrodotoxin-resistant voltage-gated sodium channel expressed by sensory neurons.
#cross-references MUID:96138382
#accession S68453
#status nucleic acid sequence not shown
#molecule_type mRNA

```

```

#residues 1-1957 #label AKO
#cross-references GB:X92184; NID:g1209466; PIDN:CAA63095.1;
#cross-references PID:e205471; PID:g1209467
#experimental_source dorsal root ganglia
CLASSIFICATION #superfamily sodium channel protein
KEYWORDS sodium channel; transmembrane protein; voltage-gated ion
channel
FEATURE
132-148 #domain transmembrane #status predicted #label TM1
158-174 #domain transmembrane #status predicted #label TM2
225-241 #domain transmembrane #status predicted #label TM3
249-265 #domain transmembrane #status predicted #label TM4
376-392 #domain transmembrane #status predicted #label TM5
666-682 #domain transmembrane #status predicted #label TM6
702-718 #domain transmembrane #status predicted #label TM7
731-747 #domain transmembrane #status predicted #label TM8
788-804 #domain transmembrane #status predicted #label TM9
865-881 #domain transmembrane #status predicted #label TM10
1156-1172 #domain transmembrane #status predicted #label TM11
1194-1210 #domain transmembrane #status predicted #label TM12
1221-1237 #domain transmembrane #status predicted #label TM13
1286-1302 #domain transmembrane #status predicted #label TM14
1400-1416 #domain transmembrane #status predicted #label TM15
1482-1498 #domain transmembrane #status predicted #label TM16
1516-1532 #domain transmembrane #status predicted #label TM17
1546-1562 #domain transmembrane #status predicted #label TM18
1606-1622 #domain transmembrane #status predicted #label TM19
1708-1724 #domain transmembrane #status predicted #label TM20
#length 1957 #molecular-weight 219913 #checksum 2137
SUMMARY
Query Match 82.0%; Score 41; DB 2; Length 1957;
Best Local Similarity 85.7%; Pred. No. 5.14e+00;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 277 IRNGTDP 283
QY 99 IRNGTHP 105
RESULT 9
ENTRY #type complete
TITLE hypothetical 8.9K protein - Rhizobium leguminosarum
ORGANISM #formal_name Rhizobium leguminosarum
DATE 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
20-Mar-1998
ACCESSION JQ0312
REFERENCE JQ0312
#authors Hontelez, J.G.J.; Lankhorst, R.K.; Katinakis, P.; van den
Bos, R.C.; van Kammen, A.
#journal Mol. Gen. Genet. (1989) 218:536-544
#title Characterization and nucleotide sequence of a novel gene fixW
upstream of the fixABC operon in Rhizobium leguminosarum.
#cross-references MUID:90066358
#accession JQ0312
#molecule_type DNA
#residues 1-79 #label HON
#cross-references GB:X16521; NID:g46184; PID:g46186
#note the authors translated the codon TCA for residue 24 as Gly
#note the gene encoding this protein is located upstream of
gene fixW
SUMMARY
#length 79 #molecular-weight 8865 #checksum 7887
Query Match 80.0%; Score 40; DB 2; Length 79;
Best Local Similarity 85.7%; Pred. No. 9.36e+00;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 28 IRNVTHP 34
QY 99 IRNGTHP 105
RESULT 10

```

```

ENTRY
TITLE      #type complete
ORGANISM   cytokine-inducible protein CIS - mouse
DATE       #formal_name Mus musculus #common_name house mouse
           27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change
           28-May-1999
ACCESSIONS S5551
REFERENCE   S5551
#authors   Yoshimura, A.; Ohkubo, T.; Kiguchi, T.; Jenkins, N.A.;
           Gilbert, D.J.; Copeland, N.G.; Hara, T.; Miyajima, A.
#journal   EMBO J. (1995) 14:2816-2826
#title     A novel cytokine-inducible gene CIS encodes an SH2-containing
           protein that binds to tyrosine-phosphorylated interleukin 3
           and erythropoietin receptors.
#cross-references MUID:95317300
#accession S5551
#status    preliminary
#molecule_type mRNA
#residues  1-257 #label YOS
#cross-references GB:D31943; NID:g1041128; PIDN:BAA06713.1;
           PID:d1007285; PID:g1041129
CLASSIFICATION #superfamily cytokine-inducible protein CIS; SH2 homology
FEATURE
82-178      #domain SH2 homology #label SH2
SUMMARY     #length 257 #molecular-weight 28536 #checksum 2728
Query Match 80.0%; Score 40; DB 2; Length 257;
Best Local Similarity 57.1%; Pred. No. 9.36e+00;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 106 VRDSTHP 112
QY 99 IRNGTHP 105

RESULT 11
ENTRY
TITLE      #type fragment
ORGANISM   zona pellucida glycoprotein C - rabbit (fragment)
           #formal_name Oryctolagus cuniculus #common_name domestic
           rabbit
DATE       28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change
           20-Aug-1999
ACCESSIONS S70401
REFERENCE   S70396
#authors   Harris, J.D.; Hibler, D.W.; Fontenot, G.K.; Hsu, K.T.;
           Yurewicz, E.C.; Sacco, A.G.
#journal   DNA Seq. (1994) 4:361-393
#title     Cloning and characterization of zona pellucida genes and
           cDNAs from a variety of mammalian species: the ZPA, ZPB and
           ZPC gene families.
#cross-references MUID:95143578
#accession S70401
#status    preliminary
#molecule_type mRNA
#residues  1-415 #label HAR
#cross-references EMBL:U05782; NID:g458280; PIDN:AAA74392.1;
           PID:g458281
CLASSIFICATION #superfamily sperm-binding glycoprotein ZP3; ZP domain
           homology
FEATURE
41-295      #domain ZP domain homology #label ZPH
SUMMARY     #length 415 #checksum 9740
Query Match 80.0%; Score 40; DB 2; Length 415;
Best Local Similarity 57.1%; Pred. No. 9.36e+00;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 196 VRTGSHP 202
QY 99 IRNGTHP 105

RESULT 12
ENTRY
TITLE      #type complete

```

```

TITLE      rnfC protein - Rhodobacter capsulatus
ORGANISM   #formal_name Rhodobacter capsulatus
DATE       27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change
           12-Feb-1999 #
ACCESSIONS S39893
REFERENCE   S39892
#authors   Schmehl, M.; Jahn, A.; Meyer zu Vilsendorf, A.; Hennecke, S.;
           Masepohl, B.; Schuppler, M.; Marxer, M.; Oelze, J.; Klipp,
           W.
#journal   Mol. Gen. Genet. (1993) 241:602-615
#title     Identification of a new class of nitrogen fixation genes in
           Rhodobacter capsulatus: a putative membrane complex
           involved in electron transport to nitrogenase.
#cross-references MUID:94088454
#accession S39893
#molecule_type DNA
#residues  1-519 #label SCH
#cross-references EMBL:X72888; NID:g435523; PID:g435525
GENETICS
#gene      rnfC
CLASSIFICATION #superfamily ferredoxin 2[4Fe-4S] homology
KEYWORDS    iron-sulfur protein
FEATURE
374-438     #domain ferredoxin 2[4Fe-4S] homology #label FER
SUMMARY     #length 519 #molecular-weight 55587 #checksum 7831
Query Match 80.0%; Score 40; DB 2; Length 519;
Best Local Similarity 71.4%; Pred. No. 9.36e+00;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 18 IRGGIHP 24
QY 99 IRNGTHP 105

RESULT 13
ENTRY
TITLE      #type complete
ORGANISM   nine-cis-epoxycarotenoid dioxygenase homolog F28J12.10 -
           Arabidopsis thaliana
ALTERNATE_NAMES hypothetical protein T9A21.200
ORGANISM     #formal_name Arabidopsis thaliana #common_name mouse-ear
           cress
DATE       23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change
           11-Jun-1999
ACCESSIONS T04531; T04937
REFERENCE   Z15377
#authors   Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.;
           Duesterhoeft, A.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.;
           Schueller, C.
#submission submitted to the Protein Sequence Database, February 1998
#accession T04531
#molecule_type DNA
#residues  1-583 #label BEV
#cross-references EMBL:AL021710
#experimental_source cultivar Columbia; BAC clone F28J12
REFERENCE   Z15390
#authors   Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.;
           Mewes, H.W.; Mayer, K.F.X.; Schueller, C.
#submission submitted to the Protein Sequence Database, February 1999
#accession T04937
#molecule_type DNA
#residues  1-377 #label BEW
#cross-references EMBL:AL021713
#experimental_source cultivar Columbia; BAC clone T9A21
GENETICS
#map_position 4
#note        F28J12.10; T9A21.200
SUMMARY     #length 583 #molecular-weight 65066 #checksum 6772
Query Match 80.0%; Score 40; DB 2; Length 583;
Best Local Similarity 71.4%; Pred. No. 9.36e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

US-09-376-430-2-14.rpr

Thu May 11 06:49:56 2000

```

Db 137 IRNGANP 143
QY 99 IRNGTHP 105

RESULT 14
ENTRY
TITLE
ORGANISM
#variety
DATE
ACCESSIONS
REFERENCE
#authors
#submission
#description
#accession
#molecule_type DNA
#residues
#cross-references EMBL:X72702
#experimental_source subtype ayw, isolate patient C1005
CLASSIFICATION #superfamily hepatitis virus DNA-directed DNA polymerase
KEYWORDS DNA biosynthesis; nucleotidyltransferase
SUMMARY #length 832 #molecular-weight 93863 #checksum 9116

Query Match 80.0%; Score 40; DB 2; Length 832;
Best Local Similarity 71.4%; Pred. NO. 9.36e+00;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 232 IRAGIHP 238
QY 99 IRNGTHP 105

RESULT 15
ENTRY
TITLE
ORGANISM
#note
DATE
ACCESSIONS
REFERENCE
#authors
#journal
#title
#cross-references MUID:85204397
#accession
#molecule_type DNA
#residues
CLASSIFICATION #superfamily hepatitis virus DNA-directed DNA polymerase
KEYWORDS DNA biosynthesis; nucleotidyltransferase
SUMMARY #length 832 #molecular-weight 93838 #checksum 9500

Query Match 80.0%; Score 40; DB 1; Length 832;
Best Local Similarity 71.4%; Pred. NO. 9.36e+00;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 232 IRAGIHP 238
QY 99 IRNGTHP 105

```

Search completed: Wed May 10 12:58:38 2000  
Job time : 8 secs.





\*\*\*\*\*  
 WIREH  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 12:52:21 2000; MasPar time 88.28 Seconds  
 2.415 Million cell updates/sec

Tabular output not generated.

Title: >US-09-376-430-2  
 Description: (99-105) from US09376430A.pap (14 of 25)  
 Perfect Score: 50  
 Sequence: 1 IRNGTHP 7

Scoring table: PAM 150  
 Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: swiss-prot38  
 1:swissprot

Statistics: Mean 19.841; Variance 16.470; scale 1.205

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	45	90.0	819	1	SYL_CHLTR LEUCYL-TRNA SYNTHETASE	1.11e+01
2	45	90.0	820	1	SYL_CHLTR LEUCYL-TRNA SYNTHETASE	1.11e+01
3	43	86.0	97	1	TEV3_BPT4 INTRON-ASSOCIATED ENDO	4.50e+01
4	41	82.0	937	1	SYL_METH LEUCYL-TRNA SYNTHETASE	1.74e+00
5	40	80.0	79	1	YFX2_RHILE HYPOTHETICAL 8.8 KD PR	3.37e+00
6	40	80.0	415	1	ZP3_RABIT ZONA PELLUCIDA SPERM-B	3.37e+00
7	40	80.0	832	1	DPOH_HPBVY DNA POLYMERASE (EC 2.7	3.37e+00
8	40	80.0	832	1	DPOH_HPBVY DNA POLYMERASE (EC 2.7	3.37e+00
9	39	78.0	365	1	DIF6_MOUSE DIFF6 PROTEIN	6.42e+00
10	39	78.0	393	1	KRL_VZVD SERINE/THREONINE-PROTE	6.42e+00
11	39	78.0	427	1	KPRL_YEAST RIBOSE-PHOSPHATE PYROP	6.42e+00
12	38	76.0	84	1	IG2R_HUMAN PUTATIVE INSULIN-LIKE	1.21e+01
13	38	76.0	155	1	RS19_DROME 40S RIBOSOMAL PROTEIN	1.21e+01
14	38	76.0	211	1	KCSA_ERWAM COLANIC ACID CAPSULAR	1.21e+01
15	38	76.0	242	1	NIFY_AZOV NIFY PROTEIN	1.21e+01
16	38	76.0	287	1	MCES_VACCC MRNA CAPPING ENZYME, S	1.21e+01
17	38	76.0	287	1	MCES_VACCV MRNA CAPPING ENZYME, S	1.21e+01
18	38	76.0	287	1	MCES_VARY MRNA CAPPING ENZYME, S	1.21e+01
19	38	76.0	413	1	AMY3_WHEAT ALPHA-AMYLASE AMY3 PRE	1.21e+01
20	38	76.0	437	1	AM3C_ORYSA ALPHA-AMYLASE ISOZYME	1.21e+01
21	38	76.0	479	1	SH2B_RAT 5-HYDROXYTRYPTAMINE 2B	1.21e+01
22	38	76.0	481	1	SH2B_HUMAN 5-HYDROXYTRYPTAMINE 2B	1.21e+01
23	38	76.0	504	1	SH2B_MOUSE 5-HYDROXYTRYPTAMINE 2B	1.21e+01

24	38	76.0	522	1	HEX1_ENTHI BETA-HEXOSAMINIDASE (E	1.21e+01
25	38	76.0	795	1	P5CS_HUMAN DELTA 1-PYRROLINE-5-CA	1.21e+01
26	38	76.0	795	1	P5CS_MOUSE DELTA 1-PYRROLINE-5-CA	1.21e+01
27	38	76.0	811	1	MUTS_THEAO DNA MISMATCH REPAIR PR	1.21e+01
28	38	76.0	818	1	MUTS_THETH DNA MISMATCH REPAIR PR	1.21e+01
29	38	76.0	855	1	MUTS_AZOV DNA MISMATCH REPAIR PR	1.21e+01
30	38	76.0	861	1	MUTS_HASIN DNA MISMATCH REPAIR PR	1.21e+01
31	38	76.0	867	1	RPO_CARMV PROBABLE RNA-DIRECTED	1.21e+01
32	38	76.0	1091	1	MSH3_MOUSE DNA MISMATCH REPAIR PR	1.21e+01
33	38	76.0	1137	1	MSH3_HUMAN DNA MISMATCH REPAIR PR	1.21e+01
34	38	76.0	2717	1	ZEPL_HUMAN ZINC FINGER PROTEIN 40	1.21e+01
35	37	74.0	159	1	NIFX_RHOCA NIFX PROTEIN	2.24e+01
36	37	74.0	191	1	SPR6_YEAST SPOULATION-SPECIFIC P	2.24e+01
37	37	74.0	257	1	YBGL_HABIN HYPOTHETICAL PROTEIN H	2.24e+01
38	37	74.0	346	1	BPHI_PSESP 4-HYDROXY-2-OXOVALERAT	2.24e+01
39	37	74.0	398	1	YJ77_CABEL HYPOTHETICAL 45.2 KD P	2.24e+01
40	37	74.0	552	1	YBJD_ECOLI HYPOTHETICAL 63.6 KD P	2.24e+01
41	37	74.0	565	1	VNUC_INCCA NUCLEOPROTEIN	2.24e+01
42	37	74.0	568	1	YOY9_CAEEL PUTATIVE UBIQUINONE BI	2.24e+01
43	37	74.0	610	1	LIN9_CAEEL LIN-9 PROTEIN	2.24e+01
44	37	74.0	642	1	LIN2_CHICK LIM DOMAIN KINASE 2 (E	2.24e+01
45	37	74.0	1431	1	TOP2_SCHPO DNA TOPOISOMERASE II (	2.24e+01

## ALIGNMENTS

RESULT	1	SYL_CHLTR	STANDARD; #	PRT; 819 AA.
AC	O84211			
DT	15-FEB-2000	(Rel. 39, Created)		
DT	15-FEB-2000	(Rel. 39, Last sequence update)		
DT	15-FEB-2000	(Rel. 39, Last annotation update)		
DE	LEUCYL-TRNA SYNTHETASE (EC 6.1.1.4)	(LEUCINE-TRNA LIGASE) (LEURS)		
GN	LEUS OR CT209.			
OS	Chlamydia trachomatis.			
OC	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=D/UV-3/CX;			
RA	MEDLINE: 99000809.			
RA	Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.;			
RA	"Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis".			
RT	Science 282:754-759(1998).			
CC	!- CATALYTIC ACTIVITY: ATP + L-LEUCINE + TRNA(LEU) = AMP + PYROPHOSPHATE + L-LEUCYL-TRNA(LEU).			
CC	!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).			
CC	!- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

EMBL; A5001294; AAC67801.1;	
DR PROSITE; P500178; AA-TRNA_LIGASE_I; 1.	
DR PFAM; PF00133; trna-synt_1; 1.	
DR PRINTS; PR00985; TRNASYNTHLEU.	
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.	
FT SIMILAR 40 50	"HIGH" REGION.
FT SIMILAR 600 603	ATP (BY SIMILARITY).
FT BINDING 603 604	"KMSKS" REGION.
FT SEQUENCE 819 AA; 92903 MW; 82412A659AFAFEFC CRC64;	

Query Match 90.08; Score 45; DB 1; Length 819;  
 Best Local Similarity 85.7%; Pred. No. 1.11e+01; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 1;

```

Db 90 IRTGTHP 96
Qy 99 IRTGTHP 105

RESULT 2
ID SYL-CHLPN STANDARD; PRT; 820 AA.
AC Q92930;
DT 15-FEB-2000 (Rel. 39, Created)
DT 15-FEB-2000 (Rel. 39, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE LEUCYL-TRNA SYNTHETASE (EC 6.1.1.4) (LEUCINE--TRNA LIGASE) (LEURS).
GN LEUS OR CPN0153.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CWL029;
RX MEDLINE; 99206606.
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
CC -!- CATALYTIC ACTIVITY: ATP + L-LEUCINE + TRNA(LEU) -> AMP +
CC PYROPHOSPHATE + L-LEUCYL-TRNA(LEU).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF001602; RAD18306.1;
CC PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
CC Aminoacyl-TRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
CC SIMILAR 40 50 "HIGH" REGION.
CC FT SIMILAR 601 605 "KMSKS" REGION.
CC FT BINDING 604 604 ATP (BY SIMILARITY).
CC SEQUENCE 820 AA; 93965 MW; 520369FC098F1926 CRC64;

Query Match 90.0%; Score 45; DB 1; Length 820;
Best Local Similarity 85.7%; Pred. No. 1,11e-01;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 90 IRTGTHP 96
Qy 99 IRTGTHP 105

RESULT 3
ID TEV3_BPT4 STANDARD; PRT; 97 AA.
AC P39514;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE INTRON-ASSOCIATED ENDONUCLEASE 3 (EC 3.1.-.-).
GN I-TEVII.
OS Bacteriophage T4.
OC viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;
OC T4-like phages.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87004574.
RA Stoeberg B.-M., Hanne S., Mathews C.Z., Mathews C.K., Rand K.N.,
RA Gait M.J.;
RT "The bacteriophage T4 gene for the small subunit of ribonucleotide
RT reductase contains an intron.";
RL EMBO J. 5:2031-2036(1986).
CC -!- FUNCTION: THIS ENDONUCLEASE IS SPECIFIC TO THE NRDB GENE SPLICE

```

```

CC JUNCTION AND IS INVOLVED IN INTRON HOMING (BY SIMILARITY).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X04140; CAA27758.1;
CC EMBL; M24598; AAA32534.1;
CC PIR; S08602; S08602.
CC Hydrolyase; Nuclease; Endonuclease; Intron homing.
CC SEQUENCE 97 AA; 11330 MW; 9CD08059375DDA91 CRC64;

Query Match 86.0%; Score 43; DB 1; Length 97;
Best Local Similarity 71.4%; Pred. No. 4.50e-01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 26 VKNGTHP 32
Qy 99 IRTGTHP 105

RESULT 4
ID SYL-METH STANDARD; PRT; 937 AA.
AC O27552;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE LEUCYL-TRNA SYNTHETASE (EC 6.1.1.4) (LEUCINE--TRNA LIGASE) (LEURS).
GN LEUS OR MH1508.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DELTA H;
RX MEDLINE; 98037514.
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noll J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- CATALYTIC ACTIVITY: ATP + L-LEUCINE + TRNA(LEU) -> AMP +
CC PYROPHOSPHATE + L-LEUCYL-TRNA(LEU).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000911; AAB85983.1;
CC PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
CC PRAM; PF00133; TRNA-synt_1; 1.
CC Aminoacyl-TRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
CC FT SIMILAR 34 44 "HIGH" REGION.
CC FT SIMILAR 609 613 "KMSKS" REGION.
CC SEQUENCE 937 AA; 108031 MW; 7EF6344EC2DC0B76 CRC64;

Query Match 82.0%; Score 41; DB 1; Length 937;
Best Local Similarity 57.1%; Pred. No. 1.74e+00;

```

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 165 VRKGHP 171  
:1:1:11  
QY 99 IRNGTHP 105

RESULT 5  
ID YFX2\_RHILE STANDARD; PRT; 79 AA.  
AC P14311;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DE HYPOTHETICAL 8.8 KD PROTEIN IN FIXW 5'REGION.  
OS Rhizobium leguminosarum.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE; 90066358.  
RA Hontelez J.G.J., Lankhorst R.K., Katinakis P., van den Bos R.C.,  
RA van Kammen A.;  
RT "Characterization and nucleotide sequence of a novel gene fixW  
RT upstream of the fixABC operon in Rhizobium leguminosarum.";  
RL Mol. Gen. Genet. 218:536-544(1989).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X16521; CAA34526.1; -  
DR PIR; JQ0312; JQ0312.  
KW Hypothetical protein.  
SQ SEQUENCE 79 AA; 8865 MW; 76A802D14B68FA39 CRC64;  
-----  
Query Match 80.0%; Score 40; DB 1; Length 79;  
Best Local Similarity 85.7%; Pred. No. 3.37e+00;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 28 IRNVTHP 34  
:1:1:111  
QY 99 IRNGTHP 105

RESULT 6  
ID ZP3\_RABIT STANDARD; PRT; 415 AA.  
AC P48833;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA  
DE GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C)  
DE (FRAGMENT).  
DE ZP3 OR ZPC.  
GN Oryctolagus cuniculus (Rabbit).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
[1]  
RN SEQUENCE FROM N.A.  
RX TISSUE-OVARY;  
RC MEDLINE; 95143578.  
RA Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,  
RA Sacco A.G.;  
RT "Cloning and characterization of zona pellucida genes and cDNAs from  
RT a variety of mammalian species: the ZPA, ZPB and ZPC gene families.";  
RL DNA Seq. 4:3361-3393(1994).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; U05782; AAA74392.1; -  
DR PROSITE; PS00682; ZP\_DOMAIN; 1.  
DR PFAM; PF00100; zona\_pellucida; 1.  
KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;  
KW Extracellular matrix; Multigene family.  
FT NON\_TER <1 18  
FT SIGNAL 1 18  
FT CHAIN 19 415  
FT DOMAIN 19 378  
FT TRANSMEM 379 399  
FT DOMAIN 400 415  
FT DOMAIN 41 301  
FT DOMAIN 415 44987 MW; 77396CF1BAA3F5CB CRC64;  
SQ SEQUENCE 415 AA; 44987 MW; 77396CF1BAA3F5CB CRC64;  
-----  
Query Match 80.0%; Score 40; DB 1; Length 415;  
Best Local Similarity 57.1%; Pred. No. 3.37e+00;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 196 VRTGSHP 202  
:1:1:11  
QY 99 IRNGTHP 105

RESULT 7  
ID DPOM\_HPBYV STANDARD; PRT; 832 AA.  
AC P04484;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE DNA POLYMERASE (EC 2.7.7.7) (VERSION 2).  
GN P.  
OS Hepatitis B virus (subtype ayw).  
OC Viruses; Retroviruses; Hepadnaviridae; Orthohepadnavirus.  
[1]  
RN SEQUENCE FROM N.A. (CLONE PHB320).  
RX MEDLINE; 85204397.  
RA Bichko V., Pushko P., Dreilina D., Pumpen P., Gren E.;  
RT "Subtype ayw variant of hepatitis B virus. DNA primary structure  
RT analysis.";  
RL FEBS Lett. 185:208-212(1985).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X02496; -; NOT\_ANNOTATED\_CDS.  
DR PIR; A00703; JDVLVB.  
DR PFAM; PF00078; rvt; 1.  
DR PFAM; PF00242; DNA\_pol\_viral\_N; 1.  
DR PFAM; PF00336; DNA\_pol\_viral\_C; 1.  
KW Transferase; DNA-directed DNA polymerase; DNA replication.  
SQ SEQUENCE 832 AA; 93838 MW; 816B8A43D6491C2C CRC64;  
-----  
Query Match 80.0%; Score 40; DB 1; Length 832;  
Best Local Similarity 71.4%; Pred. No. 3.37e+00;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 232 IRAGTHP 238  
:1:1:11  
QY 99 IRNGTHP 105

RESULT 8

ID DPOL\_HPBVA STANDARD; PRT; 832 AA.  
AC P24024;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE DNA POLYMERASE (EC 2.7.7.7).  
GN P.  
OS Hepatitis B virus (strain alpha).  
OC Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90266476.  
RA Tong S., Li J., Vitvitski L., Trepo C.;  
RT "Active hepatitis B virus replication in the presence of anti-HBe is  
associated with viral variants containing an inactive pre-C region.";  
RL Virology 176:596-603(1990).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M32138; AAA45503.1; -  
DR PIR; C34773; JVDVLA1.  
DR PFAM; PF00078; rvt; 1.  
DR PFAM; PF00242; DNA\_pol\_viral\_N; 1.  
DR PFAM; PF00336; DNA\_pol\_viral\_C; 1.  
DR Transferase; DNA-directed DNA polymerase; DNA replication.  
KW SEQUENCE 832 AA; 93589 MW; 236BA43CCADD0829 CRC64;  
SQ SEQUENCE  
Query Match 80.0%; Score 40; DB 1; Length 832;  
Best Local Similarity 71.4%; Pred. No. 3.37e+00;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Db 232 IRAGIHP 238  
|||  
QY 99 IRNGTHP 105  
-----  
RESULT 9  
ID DIF6\_MOUSE STANDARD; PRT; 365 AA.  
AC P42209;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE DIF6 PROTEIN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C57BL;  
RX MEDLINE; 91065542.  
RA Nottenburg C., Gallatin W.M., St John T.;  
RT "Lymphocyte HEV adhesion variants differ in the expression of  
multiple gene sequences.";  
RL Gene 95:279-284(1990).  
CC -!- FUNCTION: INVOLVED IN CYTOKINESIS (POTENTIAL).  
CC -!- SUBUNIT: MAY ASSEMBLE INTO A MULTICOMPONENT STRUCTURE.  
CC -!- SIMILARITY: BELONGS TO THE CDC3/CDC10/CDC11/CDC12 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M37030; AAA37803.1; -  
DR

DR PIR; JU0319; JU0319.  
DR PFAM; PF00735; GTP\_CDC; 1.  
KW Cell division; GTP-binding.  
FT NP\_BIND 32 39  
SQ SEQUENCE 365 AA; 41894 MW; 43A210D2E0A82739 CRC64;  
GTP (POTENTIAL).  
Query Match 78.0%; Score 39; DB 1; Length 365;  
Best Local Similarity 57.1%; Pred. No. 6.42e+00;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
Db 233 VRDGRP 239  
:||||:  
QY 99 IRNGTHP 105  
-----  
RESULT 10  
ID KRL\_VZVD STANDARD; PRT; 393 AA.  
AC P09251;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE SERINE/THREONINE-PROTEIN KINASE (EC 2.7.1.-).  
GN 66.  
OS Varicella-zoster virus (strain Dumas) (VZV).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 86306657.  
RA Davison A.J., Scott J.E.;  
RT "The complete DNA sequence of varicella-zoster virus.";  
RL J. Gen. Virol. 67:1759-1816(1986).  
RN [2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE; 84131932.  
RA Davison A.J.;  
RT "DNA sequence of the US component of the varicella-zoster virus  
genome.";  
RL EMBO J. 2:2203-2209(1983).  
RN [3]  
RP REVIEW.  
RX MEDLINE; 89369695.  
RA Leader D.P., Purves F.C.;  
RT "The herpesvirus protein kinase: a new departure in protein  
phosphorylation?";  
RL Trends Biochem. Sci. 13:244-246(1988).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X04370; CAA27949.1; -  
DR EMBL; X00208; CAA25031.1; -  
DR PIR; E27345; TVB566.  
DR HSP; Q16539; 1WFC.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PFAM; PF00069; pkinaase; 1.  
KW Transferase; serine/threonine-protein kinase; ATP-binding.  
FT DOMAIN 93 378  
FT NP\_BIND 99 107  
FT BINDING 122 122  
FT ACT\_SITE 206 206  
FT BY SIMILARITY.  
SQ SEQUENCE 393 AA; 43679 MW; 2396280DC40AFBF7 CRC64;  
BY SIMILARITY.  
Query Match 78.0%; Score 39; DB 1; Length 393;  
Best Local Similarity 83.3%; Pred. No. 6.42e+00;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DR PIR; S45593; S45593.  
DR SGD; L0001494; PRS1.  
DR PROSITE; PS00114; PRPP\_SYNTHETASE; 1.  
DR PFAM; PF00492; PRPP\_synt; 2.  
DR Nucleotide biosynthesis; Transferase; Kinase; Magnesium;  
KW Multigene family.  
FT METAL 128 128 MAGNESIUM (POTENTIAL).  
FT METAL 130 130 MAGNESIUM (POTENTIAL).  
FT METAL 143 143 MAGNESIUM (POTENTIAL).  
SQ SEQUENCE 427 AA; 47047 MW; P3E93ACB2E703FF CRC64;  
Query Match 78.0%; Score 39; DB 1; Length 427;  
Best Local Similarity 71.4%; Pred. No. 6.42e+00;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Db 227 IRQDHP 233  
QY 99 IRNGTHP 105  
RESULT 12  
ID IG2R\_HUMAN STANDARD; PRT; 84 AA.  
AC P09565;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DE 01-MAR-1989 (Rel. 10, Last annotation update)  
DE PUTATIVE INSULIN-LIKE GROWTH FACTOR II ASSOCIATED PROTEIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-PLACENTA;  
RX MEDLINE; 95000779.  
RA de Pagter-Holthuisen P., van der Kammen R.A., Jansen M.,  
RA van Schaik F.M.A., Sussenbach J.S.;  
RT "Differential expression of the human insulin-like growth factor II  
gene. Characterization of the IGF-II mRNAs and an mRNA encoding a  
putative IGF-II-associated protein".  
RL Biochim. Biophys. Acta 950:282-295(1988).  
CC -!- FUNCTION: NOT KNOWN.  
CC -!- MISCELLANEOUS: THE COORDINATED EXPRESSION OF INSULIN-LIKE GROWTH  
CC FACTOR (IGF II) ASSOCIATED PROTEIN AND IGF II IS DEVELOPMENTALLY  
CC AND TISSUE-SPECIFICALLY REGULATED.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X07869; CAA30710.1; -.  
DR PIR; S03384; S03384.  
KW Growth factor.  
SQ SEQUENCE 84 AA; 9081 MW; 8BFBB4E49EE5EFAB CRC64;  
Query Match 76.0%; Score 38; DB 1; Length 84;  
Best Local Similarity 71.4%; Pred. No. 1.21e+01;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Db 61 IRQDHP-67  
QY 99 IRNGTHP 105  
RESULT 13  
ID RS19\_DROME STANDARD; PRT; 155 AA.  
AC P39018;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)

Db 304 RSGTHP 309  
QY 100 RRGTHP 105  
RESULT 11  
ID KPR1\_YEAST STANDARD; PRT; 427 AA.  
AC P32895;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DE 01-NOV-1997 (Rel. 35, Last annotation update)  
DE RIBOSE-PHOSPHATE PHOSPHOKINASE 1 (EC 2.7.6.1) (PHOSPHORIBOSYL  
DE PYROPHOSPHATE SYNTHETASE 1).  
DE PRPS1 OR PRPS1 OR PRPS OR PRPS1 OR PRPS1 OR YKL181W.  
GN Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;  
OC Saccharomycetaceae; Saccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-X2180-1A;  
RX MEDLINE; 95084630.  
RA Carter A.T., Narbad A., Pearson B.M., Beck K.-F., Logghe M.,  
RA Contreras R., Schweizer M.;  
RT "Phosphoribosylpyrophosphate synthetase (PRS): a new gene family in  
RT Saccharomycetes cerevisiae".  
RL Yeast 10:1031-1044(1994).  
RN [2]  
RP ERRATUM.  
RA Carter A.T., Narbad A., Pearson B.M., Beck K.-F., Logghe M.,  
RA Contreras R., Schweizer M.;  
RL Yeast 11:191-191(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94277073.  
RA Blacketer M.J., Madaule P., Myers A.M.;  
RT "The Saccharomyces cerevisiae mutation elm4-1 facilitates  
RT pseudophthal differentiation and interacts with a deficiency in  
RT phosphoribosylpyrophosphate synthase activity to cause constitutive  
RT pseudophthal growth".  
RL Mol. Cell. Biol. 14:4671-4681(1994).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94205264.  
RA Wiemann S., Voss H., Schwager C., Rupp T., Stegemann J.,  
RA Zimmermann J., Grothues D., Sensen C., Erfle H., Hewitt N.,  
RA Bankevi A., Ansoorge W.;  
RT "Sequencing and analysis of 51.6 kilobases on the left arm of  
RT chromosome XI from Saccharomyces cerevisiae reveals 23 open reading  
RT frames including the FAS1 gene".  
RL Yeast 9:1343-1348(1993).  
CC -!- CATALYTIC ACTIVITY: ATP + D-RIBOSE 5-PHOSPHATE = AMP +  
CC -!- 5-PHOSPHO-ALPHA-D-RIBOSE 1-DIPHOSPHATE.  
CC -!- PATHWAY: THIS ENZYME IS UTILIZED BY BOTH THE DE NOVO & THE SALVAGE  
CC PATHWAYS BY WHICH ENOGENOUSLY FORMED OR EXOGENOUSLY ADDED  
CC PYRIMIDINE, PURINE, OR PYRIMIDINE BASES ARE CONVERTED TO THE  
CC CORRESPONDING RIBONUCLEOSIDE MONOPHOSPHATES.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSE-PHOSPHATE PYROPHOSPHOKINASE  
CC FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X70069; CAA49674.1; -.  
DR EMBL; L04130; AAB21811.1; -.  
DR EMBL; X74151; CAA52257.1; -.  
DR EMBL; Z28181; CAA82024.1; -.  
DR PIR; S30558; S30558.

DE 40S RIBOSOMAL PROTEIN S19.  
GN RPS19.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CANTON-S;  
RX MEDLINE; 93376521.  
RA Baumgartner S.W., Martin D., Chiquet-Ehrismann R.;  
RT "Drosophila ribosomal protein S19 cDNA sequence."  
RL Nucleic Acids Res 21:3897-3897(1993).  
CC -!- SIMILARITY: BELONGS TO THE S19E FAMILY OF RIBOSOMAL PROTEINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X73153; CAAS1677.1; -  
DR FLYBASE; FBgn010412; Rps19.  
DR PROSITE; PS00628; RIBOSOMAL\_S19E; 1.  
DR PFAM; PF01090; Ribosomal\_S19e; 1.  
KW Ribosomal protein.  
FT INIT\_MET 0  
SQ SEQUENCE 155 AA; 17174 MW; F77AB02833E7797F CRC64;

Query Match 76.0%; Score 38; DB 1; Length 155;  
Best Local Similarity 83.3%; Pred. No. 1.21e+01;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Db 83 RGVHP 88  
QY 100 RRGTHP 105  
||| |||  
||| |||

RESULT 14  
ID RCSA\_ERWAM STANDARD; PRT; 211 AA.  
AC P20098;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-MAR-1992 (Rel. 21, Last annotation update)  
DE COLANIC ACID CAPSULAR BIOSYNTHESIS ACTIVATION PROTEIN A.  
GN RCSA.  
OS Erwinia amylovora.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Erwinia.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-T;  
RX MEDLINE; 91132105.  
RA Coleman M., Pearce R., Hitchin E., Busfield F., Mansfield J.W.,  
RA Roberts I.S.;  
RT "Molecular cloning, expression and nucleotide sequence of the rcsA  
RT gene of Erwinia amylovora, encoding a positive regulator of capsule  
RT expression: evidence for a family of related capsule activator  
RT proteins."  
RL J. Gen. Microbiol. 136:1799-1806(1990).  
CC -!- FUNCTION: POSITIVE REGULATOR OF CAPSULAR POLYSACCHARIDE  
CC SYNTHESIS. RCSA AND RCSB FORM A COMPLEX TO PROMOTE TRANSCRIPTION  
CC OF THE GENES FOR CAPSULE SYNTHESIS.  
CC -!- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL  
CC REGULATORS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M57387; AAA24864.1; -  
DR PRINTS; PRO0038; HTHLUXR.  
DR PROSITE; PS00622; HTH\_LUXR\_FAMILY; 1.  
DR PFAM; PF00196; Gere; 1.  
KW Transcription regulation; DNA-binding; Activator.  
FT DNA\_BIND 159 178 H-T-H MOTIF (BY SIMILARITY).  
SQ SEQUENCE 211 AA; 24349 MW; 51BC162269735C55 CRC64;  
Query Match 76.0%; Score 38; DB 1; Length 211;  
Best Local Similarity 57.1%; Pred. No. 1.21e+01;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
Db 129 VRGTHS 135  
QY 99 IRNGTHP 105  
||| |||  
||| |||

RESULT 15  
ID NIFY\_AZQVI STANDARD; PRT; 242 AA.  
AC P14886;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-APR-1990 (Rel. 14, Last annotation update)  
DE NIFY PROTEIN.  
GN NIFY.  
OS Azotobacter vinelandii.  
OC Bacteria; Proteobacteria; gamma subdivision; Azotobacteraceae;  
OC Azotobacter.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89123097.  
RA Jacobson M.R., Brigle K.E., Bennett L.T., Setterquist R.A.,  
RA Wilson M.S., Cash V.L., Beynon J., Newton W.E., Dean D.R.;  
RT "Physical and genetic map of the major nif gene cluster from  
RT Azotobacter vinelandii."  
RL J. Bacteriol. 171:1017-1027(1989).  
CC -!- SIMILARITY: TO NIFY PROTEIN.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M20568; AAA64713.1; -  
DR PIR; B32055; B32055.  
KW Nitrogen fixation.  
SQ SEQUENCE 242 AA; 26702 MW; AC2F9DFAD6252D63 CRC64;

Query Match 76.0%; Score 38; DB 1; Length 242;  
Best Local Similarity 57.1%; Pred. No. 1.21e+01;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Db 192 VRGVHP 198  
QY 99 IRNGTHP 105  
||| |||  
||| |||

Search completed: Wed May 10 12:53:58 2000  
Job time: 97 secs.

\*\*\*\*\*  
 M P S R C H  
 \*\*\*\*\* (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 12:54:17 2000; MasPar time 225.58 Seconds  
 Tabular output not generated. 2.152 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (99-105) from US09376430A.pep (14 of 25)  
 Perfect Score: 50  
 Sequence: 1 IRNGTHP 7

Scoring table: PAM 150  
 Gap 11

Searched: 225878 segs, 69334122 residues  
 Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: sptrembl12  
 1:sp\_archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
 5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
 9:sp\_phase 10:sp\_plant 11:sp\_rodent 12:sp\_unclassified  
 13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 19.420; Variance 15.902; scale 1.221

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	45	90.0	819	2	084211	LEUCYL TRNA SYNTHETASE
2	45	90.0	820	2	092930	LEUCYL TRNA SYNTHETASE
3	45	90.0	825	14	067885	DNA POLYMERASE (EC 2.7
4	45	90.0	832	14	067878	DNA POLYMERASE (EC 2.7
5	45	90.0	832	14	081169	DNA POLYMERASE (EC 2.7
6	43	86.0	269	9	038419	NRDB.
7	43	86.0	833	11	092106	MUTS HOMOLOG.
8	42	84.0	791	14	083424	REPLICASE.
9	41	82.0	299	5	09XXA2	VM02B12L.3 PROTEIN.
10	41	82.0	1956	11	062968	SODIUM CHANNEL PN3.
11	41	82.0	1957	11	063554	VOLATAGE-GATED SODIUM
12	40	80.0	256	11	070512	CYTOKINE-INDUCIBLE SH2
13	40	80.0	257	11	062225	CYTOKINE-INDUCIBLE SH2
14	40	80.0	291	14	067929	DNA POLYMERASE (EC 2.7
15	40	80.0	291	14	067932	DNA POLYMERASE (EC 2.7
16	40	80.0	361	4	09V5R1	CYTOKINE-INDUCIBLE SH2
17	40	80.0	390	14	067908	HBSAG (FRAGMENT).
18	40	80.0	519	2	Q52716	NITROGEN FIXATION IRON
19	40	80.0	583	10	049505	NEOXANTHIN CLEAVAGE EN
20	40	80.0	832	14	Q9WNS2	POLYMERASE PRECURSOR.

21	40	80.0	832	14	Q96846	COMPLETE GENOME.	5.47e+00
22	40	80.0	832	14	067892	DNA POLYMERASE (EC 2.7	5.47e+00
23	40	80.0	832	14	067919	DNA POLYMERASE (EC 2.7	5.47e+00
24	40	80.0	832	14	067907	DNA POLYMERASE (EC 2.7	5.47e+00
25	40	80.0	832	14	067913	DNA POLYMERASE (EC 2.7	5.47e+00
26	40	80.0	843	14	09YL91	POLYMERASE.	5.47e+00
27	40	80.0	845	14	09WRK2	POLYMERASE.	5.47e+00
28	40	80.0	2209	14	09Y225	L PROTEIN.	5.47e+00
29	40	80.0	3124	14	066237	349-KDA VIRAL POLYPROT	5.47e+00
30	39	78.0	89	2	044590	PLASMOD PTOM9 FROM ALC	1.07e+01
31	39	78.0	299	14	088980	MINOR CAPSID LIKE PROT	1.07e+01
32	39	78.0	1880	14	092358	NON-STRUCTURAL POLYPRO	2.05e+01
33	38	76.0	84	14	072024	NON-STRUCTURAL PROTEIN	2.05e+01
34	38	76.0	264	12	087756	NITRIE REDUCTASE (FRA	2.05e+01
35	38	76.0	287	14	057213	MRNA CAPPING ENZYME.	2.05e+01
36	38	76.0	318	24	092530	PROBABLE RIBOFLAVIN KI	2.05e+01
37	38	76.0	377	2	069760	POTATIVE PUTRESCINE/SP	2.05e+01
38	38	76.0	439	10	041770	LEUCOCYTE DNA BINDING	2.05e+01
39	38	76.0	445	13	013038	NEOXANTHIN CLEAVAGE EN	2.05e+01
40	38	76.0	595	10	049675	STT3 PROTEIN (FRAGMENT	2.05e+01
41	38	76.0	723	5	097353	PUT. P77-80 PROTEIN (A	2.05e+01
42	38	76.0	763	14	09WJEO	(CARMV) COMPLETE GENOM	2.05e+01
43	38	76.0	869	14	09WJDO	ALPHA-LATROCRUSTOTOXIN	2.05e+01
44	38	76.0	1395	5	09XZCO	DNA-BINDING PROTEIN (M	2.05e+01
45	38	76.0	1902	4	Q14122		

## ALIGNMENTS

RESULT 1	PRELIMINARY;	PRT;	819 AA.
ID 084211			
AC 084211;1998	(TRENBLrel. 08, Created)		
DT 01-NOV-1998	(TRENBLrel. 08, Last sequence update)		
DT 01-NOV-1999	(TRENBLrel. 12, Last annotation update)		
DE LEUCYL TRNA SYNTHETASE.			
GN LEUS.			
OS Chlamydia trachomatis.			
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=D/UV-3/CX;			
RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L., MITCHELL W.P., OLINGER L., TATSOV R.L., ZHAO Q., KOONIN E.V., DAVIS R.W.;			
RA "Genome Sequence of an Obligate Intracellular Pathogen of Humans: Chlamydia trachomatis."			
RT Science 0:0-0(1998).			
EL [2]			
RP SEQUENCE FROM N.A.			
RC STRAIN=D/UV-3/CX;			
RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L., MITCHELL W.P., OLINGER L., TATSOV R.L., ZHAO Q., KOONIN E.V., DAVIS R.W.;			
RA Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.			
RL EMBL; AEO01294; AAC67801.1; -			
DR PFAM; PF00133; ARCA-synt_1; 1.			
DR PRINTS; PR00985; TRNASYNTHLEU.			
SW Aminoacyl-tRNA synthetase.			
KW SEQUENCE 819 AA; 92903 MW; 23F26818 CRC32;			
Query Match	90.0%;	Score 45;	DB 2; Length 819;
Best Local Similarity	85.7%;	Pred. No. 1.63e-01;	
Matches	6;	Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Db	90	IRNGTHP_96	
Qy	99	IRNGTHP_105	
RESULT 2	PRELIMINARY;	PRT;	820 AA.
ID Q92930			
AC Q92930;			

DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE LEUCYL TRNA SYNTHETASE.  
 GN LEUS.  
 OS Chlamydia pneumoniae.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiaophila.  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CWL029;  
 RA KALMAN S., MITCHELL W., MARATHE R., LAMMEL C., FAN J., OLINGER L.,  
 RA GRIMWOOD J., DAVIS R.W., STEPHENS R.S.;  
 RT "Comparative Genomes of Chlamydia pneumoniae and C. trachomatis."  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE001602; AAD18306.1;  
 SQ SEQUENCE 820 AA; 93965 MW; CC0333EB CRC32;  
 Query Match 90.0%; Score 45; DB 2; Length 820;  
 Best Local Similarity 85.7%; Pred. No. 1.63e-01;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 90 IRTGTHP 96  
 QY 99 IRTGTHP 105  
 RESULT 3  
 ID Q67885 PRELIMINARY; PRT; 825 AA.  
 AC Q67885;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE DNA POLYMERASE (EC 2.7.7.7).  
 GN P.  
 OS Hepatitis B virus.  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HBSAG SUBTYPE AYW;  
 RX MEDLINE; 92020153.  
 RA LAI M.E., MELIS A., MAZZOLENI A.P., UCCHEDDU P., BALESTRIERI A.;  
 RT "Sequence analysis of hepatitis B virus genome of a new mutant of ayw  
 subtype isolated in Sardinia."  
 RL Nucleic Acids Res. 19:5078-5078 (1991).  
 DR EMBL; X59795; CAA42466.1;  
 DR PFAM; PF00336; DNA\_pol\_viral\_C; 1.  
 DR PFAM; PF00242; DNA\_pol\_viral\_N; 1.  
 DR PFAM; PF00078; rvt; 1.  
 KW Transferase; DNA-directed DNA polymerase; DNA replication.  
 SQ SEQUENCE 825 AA; 93073 MW; 16B93F76 CRC32;  
 Query Match 90.0%; Score 45; DB 14; Length 825;  
 Best Local Similarity 85.7%; Pred. No. 1.63e-01;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 232 IRTGTHP 238  
 QY 99 IRTGTHP 105  
 RESULT 4  
 ID Q67878 PRELIMINARY; PRT; 832 AA.  
 AC Q67878;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE DNA POLYMERASE (EC 2.7.7.7).  
 GN P.  
 OS Hepatitis B virus.  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PATIENT CI, HBV SUBTYPE AYW;

RA LAI M.E., MAZZOLENI A.P., BALESTRIERI A., MELIS A., PORRU A.;  
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X55258; CAA46356.1;  
 DR PFAM; PF00336; DNA\_pol\_viral\_C; 1.  
 DR PFAM; PF00242; DNA\_pol\_viral\_N; 1.  
 DR PFAM; PF00078; rvt; 1.  
 KW Transferase; DNA-directed DNA polymerase; DNA replication.  
 SQ SEQUENCE 832 AA; 93871 MW; 31BBA72B CRC32;  
 Query Match 90.0%; Score 45; DB 14; Length 832;  
 Best Local Similarity 85.7%; Pred. No. 1.63e-01;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 232 IRTGTHP 238  
 QY 99 IRTGTHP 105  
 RESULT 5  
 ID Q81169 PRELIMINARY; PRT; 832 AA.  
 AC Q81169;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE DNA POLYMERASE (EC 2.7.7.7).  
 GN P.  
 OS Hepatitis B virus.  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94149856.  
 RA HASEGAWA K., HUANG J., ROGERS S.A., BLUM H.E., LIANG T.J.;  
 RT "Enhanced replication of a hepatitis B virus mutant associated with an  
 epidemic of fulminant hepatitis."  
 RL J. Virol. 68:1651-1659 (1994).  
 DR EMBL; L27106; AAA18583.1;  
 DR PFAM; PF00336; DNA\_pol\_viral\_C; 1.  
 DR PFAM; PF00242; DNA\_pol\_viral\_N; 1.  
 DR PFAM; PF00078; rvt; 1.  
 KW Transferase; DNA-directed DNA polymerase; DNA replication.  
 SQ SEQUENCE 832 AA; 93705 MW; C0593B80 CRC32;  
 Query Match 90.0%; Score 45; DB 14; Length 832;  
 Best Local Similarity 85.7%; Pred. No. 1.63e-01;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 232 IRTGTHP 238  
 QY 99 IRTGTHP 105  
 RESULT 6  
 ID Q38419 PRELIMINARY; PRT; 269 AA.  
 AC Q38419;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE NRDB.  
 GN I-TEVII.  
 OS Bacteriophage RB3.  
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;  
 OC T4-like phages.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA EDDY S.R., GOLD L.;  
 RL Genes Dev. 0:0-0 (1991).  
 DR EMBL; X59078; CAA41802.1;  
 SQ SEQUENCE 269 AA; 30814 MW; 418C8785 CRC32;  
 Query Match 86.0%; Score 43; DB 9; Length 269;  
 Best Local Similarity 71.4%; Pred. No. 6.88e-01;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;



US-09-376-430-2-14.rspt

Thu May 11 06:49:57 2000

Db 198 VRNGTHP 204  
QY 99 IRNGTHP 105

RESULT 7 PRELIMINARY; PRT; 833 AA.  
ID Q921Q6  
AC Q921Q6  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE MTS HOMOLOG.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-129.  
RA ROSEN L., MADAN A., OIN S., SHAFFER T., RATCLIFFE A., ABBASI N.,  
RA DICKHOFF R., JAMES R., LORETZ C., LASKY S., HOOD L.,  
RT "Sequence of the mouse major histocompatibility locus class III  
region." (NOV-1998) to the EMBL/GenBank/DBJ databases.  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF09905; AAC84154.1  
DR PROSITE; P500486; DNA\_MISMATCH\_REPAIR\_2; 1.  
SQ SEQUENCE 833 AA; 92574 MW; 274DC21A CRC32;

Query Match 86.0%; Score 43; DB 11; Length 833;  
Best Local Similarity 85.7%; Pred. No. 6.88e-01; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 1;  
Db 557 IRNGRHP 563  
QY 99 IRNGTHP 105

RESULT 8 PRELIMINARY; PRT; 791 AA.  
ID Q83424  
AC Q83424  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE REPLICASE.  
GN P89.  
OS Melon necrotic spot virus (MNSV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;  
OC Carmovirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90063551.  
RA RIVIERE C.J., POT J., TREMAINE J.H., ROCHON D.M.;  
RT "Coat protein of melon necrotic spot carmovirus is more similar to  
those of tombusviruses than those of carmoviruses.";  
RL J. Gen. Virol. 70:3033-3042(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91011332.  
RA RIVIERE C.J., ROCHON D.M.;  
RT "Nucleotide sequence and genomic organization of melon necrotic spot  
virus.";  
RL J. Gen. Virol. 71:1887-1896(1990).  
DR EMBL; M29671; AAB02430.1;  
DR EMBL; D12536; BAA02099.1;  
DR PFAM; PF01615; PV\_RDRG; 1.  
SQ SEQUENCE 791 AA; 86681 MW; BAD0991A CRC32;

Query Match 84.0%; Score 42; DB 14; Length 791;  
Best Local Similarity 71.4%; Pred. No. 1.39e+00; Indels 0; Gaps 0;  
Matches 5; Conservative 1; Mismatches 1;  
Db 282 VRNGTHP 288  
QY 99 IRNGTHP 105

RESULT 9 PRELIMINARY; PRT; 299 AA.

ID Q9XXA2  
AC Q9XXA2  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE VW02B12L.3 PROTEIN.  
DE VW02B12L.3  
GN VW02B12L.3  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94150718.  
RA MORTIMORE B.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans.";  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER N., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKIN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIRRY-MIES J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans.";  
RL Nature 368:32-38(1994).  
DR EMBL; AL031269; CAA20332.1;  
SQ SEQUENCE 299 AA; 33034 MW; 13B73EC7 CRC32;

Query Match 82.0%; Score 41; DB 5; Length 299;  
Best Local Similarity 71.4%; Pred. No. 2.78e+00; Indels 0; Gaps 0;  
Matches 5; Conservative 2; Mismatches 0;  
Db 182 VRNGTRP 188  
QY 99 IRNGTHP 105

RESULT 10 PRELIMINARY; PRT; 1956 AA.  
ID Q62968  
AC Q62968  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE SODIUM CHANNEL PN3  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96198040.  
RA SANGAMESWARAN L.B., DELGADO S.G., FISH L.M., KOCH B.D., JAKEMAN L.B.,  
RA STEWART G.R., SIZE P., HUNTER J.C., EGLEN R.M., HERMAN R.C.;  
RL J. Biol. Chem. 271:13292-13292(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96198040.  
RA SANGAMESWARAN L.B., DELGADO S.G., FISH L.M., KOCH B.D., JAKEMAN L.B.,  
RA STEWART G.R., SIZE P., HUNTER J.C., EGLEN R.M., HERMAN R.C.;  
RT "Structure and function of a novel voltage-gated, tetrodotoxin-  
resistant sodium channel specific to sensory neurons.";  
RL J. Biol. Chem. 271:5953-5957(1996).  
DR EMBL; U53833; AAC52619.1;  
DR PFAM; PF00520; Ion\_trans; 4.

DR PRINTS: PR00170; NACHANNEL.  
KW Ionic channel.  
SQ SEQUENCE 1956 AA; 219731 MW; 9F1710DA CRC32;

Query Match 82.0%; Score 41; DB 11; Length 1956;  
Best Local Similarity 85.7%; Pred. No. 2.78e+00;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 277 IRNGTDP 283

QY 99 IRNGTHP 105

RESULT 11  
ID Q63554 PRELIMINARY; PRT; 1957 AA.  
AC Q63554;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE VOLATAGE-GATED SODIUM CHANNEL.  
GN SNS.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RQ TISSUE-DRG;  
RX MEDLINE; 96138382.  
RA AKOPIN A.N.; SIVILOTTI L.; WOOD J.N.;  
RT "A tetrodotoxin-resistant voltage-gated sodium channel expressed by  
sensory neurons."  
RL Nature 379:257-262(1995).  
DR EMBL; X92184; CAA63095.1;  
DR PFAM; PF00520; ion.trans. 4.  
DR PRINTS; PR00170; NACHANNEL.  
KW Ionic channel.  
SQ SEQUENCE 1957 AA; 219913 MW; ABA9DB6C CRC32;

Query Match 82.0%; Score 41; DB 11; Length 1957;  
Best Local Similarity 85.7%; Pred. No. 2.78e+00;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 277 IRNGTDP 283

QY 99 IRNGTHP 105

RESULT 12  
ID Q70512 PRELIMINARY; PRT; 256 AA.  
AC Q70512;  
DT 01-AUG-1998 (TREMBlrel. 07, Created)  
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE CYTOKINE-INDUCIBLE SH2-CONTAINING PROTEIN (FRAGMENT).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RQ STRAIN-SPRAGUE-DAWLEY; TISSUE=LIVER;  
RA SMITH R.J.; FITZGIBBONS T.P.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF065161; AAC17502.1;  
DR HSSP; P16277; 1BLJ.  
DR PFAM; PF00017; SH2; 1.  
FT NON\_TER 1  
SQ SEQUENCE 256 AA; 28398 MW; 9EC5945F CRC32;

Query Match 80.0%; Score 40; DB 11; Length 256;  
Best Local Similarity 57.1%; Pred. No. 5.47e+00;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 105 VRDSTHP 111

QY 99 IRNGTHP 105

RESULT 13  
ID Q62225 PRELIMINARY; PRT; 257 AA.  
AC Q62225;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE CYTOKINE INDUCIBLE SH2-CONTAINING PROTEIN.  
GN CISH OR CIS.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RQ MEDLINE; 95317300.  
RA YOSHIMURA A.; OHKUBO T.; KIGUCHI T.; JENKINS N.A.; GILBERT D.J.;  
RA COPELAND N.G.; HARA T.; MIYAJIMA A.;  
RT "A novel cytokine-inducible gene CISH encodes an SH2-containing protein  
that binds to tyrosine-phosphorylated interleukin 3 and erythropoietin  
receptors."  
RL EMBO J. 14:2816-2826(1995).  
DR EMBL; D31943; BAA06713.1;  
DR HSSP; P16277; 1BLJ.  
DR MGD; MGI:103159; Cish.  
DR PFAM; PF00017; SH2; 1.  
SQ SEQUENCE 257 AA; 28536 MW; A0017153 CRC32;

Query Match 80.0%; Score 40; DB 11; Length 257;  
Best Local Similarity 57.1%; Pred. No. 5.47e+00;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 106 VRDSTHP 112

QY 99 IRNGTHP 105

RESULT 14  
ID Q67929 PRELIMINARY; PRT; 291 AA.  
AC Q67929;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE DNA POLYMERASE (EC 2.7.7.7) (FRAGMENT).  
GN P.  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RQ STRAIN=AYW;  
RA KARAYIANNIS P.;  
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X80244; CAA56878.1;  
DR PFAM; PF00242; DNA\_pol\_viral\_N; 1.  
KW Transferrase; DNA-directed DNA polymerase; DNA replication.  
FT NON\_TER 291  
SQ SEQUENCE 291 AA; 32975 MW; 977520F5 CRC32;

Query Match 80.0%; Score 40; DB 14; Length 291;  
Best Local Similarity 71.4%; Pred. No. 5.47e+00;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 232 IRAGIHP 238

QY 99 IRNGTHP 105

RESULT 15  
ID Q67932 PRELIMINARY; PRT; 291 AA.  
AC Q67932;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)

US-09-376-430-2-14.rspt

Thu May 11 06:49:57 2000

DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
 DE 01-NOV-1998 (TREMELrel. 08, Last annotation update)  
 GN DNA POLYMERASE (EC 2.7.7.7) (FRAGMENT).  
 OS Hepatitis B virus.  
 OC Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AIW;  
 RA KARAYIANNIS P.;  
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X80926; CAA56892.1; -;  
 KW PFAM; PF00242; DNA-pol\_viral\_N; 1.  
 DR Transferase; DNA-directed DNA polymerase; DNA replication.  
 FT NON\_TER 291  
 SQ SEQUENCE 291 AA; 32917 MW; 122C02DC CRC32;

Query Match 80.0%; Score 40; DB 14; Length 291;  
 Best Local Similarity 71.4%; Pred. No. 5.47e+00;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 232 IRAGIHP 238  
 QY 99 IRNGTHP 105

Search completed: Wed May 10 12:58:13 2000  
 Job time : 236 secs.



\*\*\*\*\*

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Run on: Wed May 10 13:07:07 2000; MasPar time 2.82 Seconds  
Tabular output not generated. 75.574 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (118-126) from US09376430A.pap (15 of 25)  
Perfect Score: 63  
Sequence: 1 KPSSPKHVR 9

Scoring table: PAM 150  
Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq35  
1:geneseqp

Statistics: Mean 15.138; Variance 39.993; scale 0.379

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	47	74.6	459	1 R42848	VIP receptor protein.	5.07e+01
2	47	74.6	897	1 R99800	NTII-1 nerve protein.	5.07e+01
3	45	71.4	515	1 W38935	Peptide resembling an	8.92e+01
4	45	71.4	1146	1 W03137	Herpesvirus of turkeys	8.92e+01
5	45	71.4	1146	1 W96322	Inducible nitric oxide	8.92e+01
6	45	71.4	1153	1 W36113	Human inducible nitric	8.92e+01
7	45	71.4	1153	1 R55764	Sequence encoded by th	8.92e+01
8	45	71.4	1153	1 R88464	Hepatocyte inducible n	8.92e+01
9	45	71.4	1153	1 R63206	Nitric-oxide-synthase	8.92e+01
10	44	69.8	43	1 W38202	Arabidopsis SCARECROW	1.18e+02
11	44	69.8	265	1 Y02653	Human secreted protein	1.18e+02
12	44	69.8	350	1 W80397	A secreted protein enc	1.18e+02
13	44	69.8	653	1 W38178	Arabidopsis SCARECROW	1.18e+02
14	44	69.8	657	1 W20051	Methionyl-tRNA synthet	1.18e+02
15	44	69.8	665	1 W27660	Streptococcus pneumoni	1.18e+02
16	44	69.8	895	1 R80231	Human dystroglycan CDN	1.18e+02
17	44	69.8	895	1 W31589	Human dystrophin-associ	1.18e+02
18	44	69.8	1203	1 W81572	Mus dunni endogenous v	1.18e+02
19	42	66.7	323	1 R39869	C peptide RV-C2, resid	2.05e+02
20	42	66.7	323	1 W97866	Human cytokine recepto	2.05e+02
21	42	66.7	350	1 W62236	Streptococcus pneumoni	2.05e+02
22	42	66.7	350	1 W34666	Partial PSD-95 protein	2.05e+02
23	42	66.7	523	1 W72229	HSV-2 strain SB5 Conti	2.05e+02

24	42	66.7	574	1 W97861	Human cytokine recepto	2.05e+02
25	42	66.7	610	1 W72228	HSV-2 strain SB5 Conti	2.05e+02
26	42	66.7	649	1 W72097	HSV-2 strain SB5 Conti	2.05e+02
27	42	66.7	882	1 W34662	Partial PSD-93 protein	2.05e+02
28	42	66.7	895	1 W31588	Rabbit (156 kDa) dystri	2.05e+02
29	42	66.7	895	1 R40929	43 kDa/156 kDa non-dys	2.05e+02
30	42	66.7	895	1 R50240	Sequence of precursor	2.05e+02
31	42	66.7	895	1 R80230	Rabbit 97 kDa dystrogl	2.05e+02
32	42	66.7	992	1 R73917	Rubella virus Therien	2.05e+02
33	42	66.7	992	1 R47257	Pre-pro-DPP	2.05e+02
34	41	65.1	950	1 W83319	Mouse SWAD interacting	2.69e+02
35	40	63.5	401	1 W32092	Porcine retrovirus par	3.53e+02
36	40	63.5	423	1 W8403	FelV-A pol protein.	3.53e+02
37	40	63.5	445	1 W25768	Human MN 64.	3.53e+02
38	40	63.5	800	1 R28608	Potential tyrosine kin	3.53e+02
39	40	63.5	800	1 W83165	Rat orphan tyrosine ki	3.53e+02
40	40	63.5	937	1 W83167	Human receptor tyrosin	3.53e+02
41	40	63.5	1135	1 W31185	Human p160 polypeptid	3.53e+02
42	40	63.5	1194	1 W39272	Porcine retrovirus POL	3.53e+02
43	40	63.5	1784	1 R94427	FelV F6A provirus clon	3.53e+02
44	40	63.5	2206	1 R22210	True type 3 poliovirus	3.53e+02
45	40	63.5	2954	1 Y01632	Amino acid sequence of	3.53e+02

## ALIGNMENTS

RESULT 1  
ID R42848 standard; Protein; 459 AA.  
AC R42848;  
DT 13-MAY-1994 (first entry)  
DE VIP receptor protein.  
KW vasoactive intestinal polypeptide receptor; VIP; rat; binding;  
KW adenylyate cyclase activity; stimulus.  
OS Rattus rattus.  
PN R0255394.A.  
PD 05-CCI-1993;  
PF 13-FEB-1992; 026607.  
PR 13-FEB-1992; JP-026607.  
PA (OSAB-) ZH OSAKA BIOSCIENCE KENKYUSHO.  
DR WPI; 93-348480/44.  
DR N-PSDB; Q50349.  
PT Vasoactive intestinal polypeptide - prep'd. in large amt. by  
culturing microbe transformed by new DNA coding polypeptide  
Claim 2; Page 6; 14pp; Japanese.  
PS The sequence can be used to produce large amounts of the VIP  
receptor peptide, by culturing a microorganism transformed by  
the sequence.  
SQ Sequence 459 AA;

Query Match 74.6%; Score 47; DB 1; Length 459;  
Best Local Similarity 66.7%; Pred. No. 5.07e+01;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 2 RPSPPHVR 10

QY 118 KPSSPKHVR 126

RESULT 2  
ID R99800 standard; Protein; 897 AA.

AC R99800;  
DT 02-JAN-1997 (first entry)  
DE NTII-1 nerve protein, facilitates regeneration of nerve cells.  
KW Probe; central nervous system; peripheral nervous system; CNS; PNS;  
axon; neurone; neuronal cells; glial cells; Schwann cells; trauma;  
pathology; regeneration; MS2; metalloprotease.  
OS Homo sapiens.  
PN WO9617865-A2.  
PD 13-JUN-1996.  
PF 05-DEC-1995; E04777.  
PR 05-DEC-1994; BE-443159.  
PR 27-JAN-1995; DE-002525.  
PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Gillen C, Gleichmann M, Mueller H;  
DR WPI; 96-287114/29.  
PT N-P5DB; T34615.  
PT Nucleic acids associated with damaged or regenerating nerve cells  
PT for diagnostic or therapeutic use.  
PS Claim 4; Page 46-49; 54pp; German.  
CC Nucleic acids described in T34587-T34616 are useful as probes for  
CC detecting traumatic and pathological changes in the peripheral  
CC nervous system. They may also be used to identify hybridising  
CC sequences from an ischial nerve cDNA library, these sequences encode  
CC proteins expressed in damaged and/or regenerating nerve cells. They  
CC can be used to transfect cells which then express the protein which  
CC can then be harvested and studied. The nucleic acids are preferably  
CC selected from NTII-11 which corresponds to plasmolipin cDNA with a  
CC 234 bp N-terminal extension and is expressed in glial cells of the  
CC central nervous system and Schwann cells of the peripheral nervous  
CC system; NTII-1, which codes for this protein having 79% homology  
CC with human CDC4L in three domains and may be involved in regulating  
CC Schwann cell growth as well as being therapeutically useful for  
CC regenerating nerve cells; and CR11-7 which codes for a protein  
CC which is similar to the macrophage-specific cysteine-rich mouse  
CC protein M2 and a metalloprotease and which inhibits interaction  
CC between Schwann cells and the extracellular matrix and is useful  
CC for the therapy of nerve damage by facilitating regeneration of  
CC damaged axons.  
SQ Sequence 887 AA;

Query Match 74.6%; Score 47; DB 1; Length 887;  
Best Local Similarity 66.7%; Pred. No. 5.07e+01;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 334 KPSPKRV 342

Qy 118 KPSPKRV 126

RESULT 3  
ID W39935 standard; peptide; 15 AA.  
AC W39935.  
DE Peptide resembling an SH3 domain binding peptide SEQ ID NO:332.  
KW Cortactin; SH3 domain; binding peptide; Src homology region 3.  
KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck;  
KW Abi; pLcgamma; p53bp2; Crk; Yes; Grb2.  
OS Synthetic.  
PN WO9730074-A1.  
PD 21-AUG-1997.  
PF 14-FEB-1997; U02298.  
PF 16-FEB-1996; US-602999.  
PA (CYTO-) CYTOGEN CORP.  
PA (UNNC-) UNIV NORTH CAROLINA.  
PI Der CJ, Fowlkes DM, Kay BK, Quilliam LA, Rider JE,  
PI Sparks AB, Thorn JM;  
DR WPI; 97-424972/39.  
PT Src homology region 3 binding peptide - used to activate Src  
PT tyrosine kinase(s) and to stimulate immune response by increasing  
PT production of certain lymphokine(s), e.g. interleukin-1  
PS Claim 22; Page 91; 131pp; English.  
CC The present sequence represents a peptide which resembles a Src homology  
CC region 3 (SH3) binding peptide. SH3 binding peptides are selected from:  
CC (a) peptides which bind the SH3 domain of Cortactin; (b) peptides which  
CC bind the middle SH3 domain of Nck; (c) peptides which bind the SH3  
CC domain of Abi; (d) peptides which bind the SH3 domain of Src; (e)  
CC peptides which bind the SH3 domain of PLC gamma; (f) peptides which bind  
CC the SH3 domain of p53bp2; (g) peptides which bind the amino-terminal SH3  
CC domain of Crk; (h) peptides which bind the SH3 domain of Yes; and (i)  
CC peptides which bind the amino-terminal SH3 domain of Grb2. The purified  
CC binding peptides can be used in the method to identify inhibitors of  
CC their binding to their respective SH3 domains, which could be used to  
CC modulate the pharmacological activity of proteins or polypeptide  
CC containing the SH3 domain. The peptides can also be used to activate  
CC Src or Src-related protein tyrosine kinases, to stimulate the immune  
CC response by increasing the production of certain lymphokines, e.g.

CC tumour necrosis factor-alpha and interleukin-1, or to deliver a  
CC conjugated molecule to certain cellular compartments containing Src or  
CC Src related proteins.  
SQ Sequence 15 AA;

Query Match 71.4%; Score 45; DB 1; Length 15;  
Best Local Similarity 44.4%; Pred. No. 8.92e+01;  
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 5 RPPPKRIR 13

Qy 118 KPSPKRV 126

RESULT 4  
ID W03137 standard; Protein; 514 AA.  
AC W03137;  
DE 18-MAR-1997 (first entry)  
DE Herpesvirus of turkeys UL13 protein.  
KW HVT; turkey herpes virus; THV; UL13; Marek; live avian vaccine;  
KW Gumboro disease; infectious bursal disease virus; IBDV.  
OS Herpesvirus of turkeys.  
PN WO9621034-A1.  
PD 11-JUL-1996.  
PF 29-DEC-1995; F01763.  
PF 30-DEC-1994; FR-016016.  
PI (INMR ) RHONE MERIEUX SA.  
PI Audonnet JF, Bublout MJM, Dartail RJ, Duinat CV;  
PI Laplace ELF, Riviere MAE;  
DR WPI; 96-334009/33.  
DR N-PSDB; T35874.  
PT Live avian vaccine based on Marek disease virus - has sequence  
PT encoding antigenic polypeptide inserted into the UL13 gene  
PS Example 6; Fig 9; 75pp; French.  
CC A 14.5 Kb BamHI D fragment of herpesvirus of turkeys strain FC126  
CC DNA was cloned in vector pBR322 to give the plasmid pRD066. The  
CC present sequence was deduced from a 5471 bp region of the fragment  
CC comprising an open reading frame with homology to the UL13 gene of  
CC herpes simplex virus HSV-1. Fragments of the UL13 ORF sequence were  
CC obtained, e.g. by PCR amplification or by restriction enzyme  
CC digestion of plasmid pRD066. The fragments were subsequently used  
CC in the construction of live avian viral vaccines in which antigens  
CC from e.g. infectious bursal disease virus or Newcastle disease  
CC virus are inserted into the UL13 ORF.  
SQ Sequence 514 AA;

Query Match 71.4%; Score 45; DB 1; Length 514;  
Best Local Similarity 85.7%; Pred. No. 8.92e+01;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 109 PSSPKHI 115

Qy 119 PSSPKHV 125

RESULT 5  
ID W96322 standard; Protein; 1146 AA.  
AC W96322;  
DE 28-JUN-1999 (first entry)  
DE Inducible nitric oxide synthase.  
KW Manganese containing superoxide dismutase; MnSOD; IDDM;  
KW diabetes mellitus; treatment; therapy; nitric oxide; NO; beta cell;  
KW fatty acid; lipotoxic; cytotoxic; cytokine; osteoporosis;  
KW inflammatory disease; autoimmune disease; neurodegenerative disease.  
OS Homo sapiens.  
PN WO9506059-A2.  
PD 11-FEB-1999.  
PF 30-JUL-1998; U15781.  
PF 03-MAR-1998; US-055092.  
PR 30-JUL-1997; US-055092.  
PA (BETA-) BETAGENE INC.  
PA (TEXA ) UNIV TEXAS SYSTEM.  
PI Clark SA, Hommeier H, Koyama K, Lee Y, Newgard CB,

PI Ohneda M, Shimabukuro M, Thigpen A, Unger RH;  
DR WPI: 99-153446/13.  
DR N-PSDB: X08434.  
PT - used for treating, e.g. diabetes, obesity, wasting syndromes,  
PT osteoporosis, inflammatory diseases, autoimmune diseases or  
PT neurodegenerative diseases  
PS Disclosure: Page 247-251; 253pp; English.  
CC Inhibition of cytokine mediated immunotoxicity of cells can be  
CC achieved by blocking free radical production or the accumulation of  
CC free radicals in that cell. Treatment of insulin dependent diabetes  
CC mellitus (IDDM) can be achieved by blocking nitric oxide (NO)  
CC production in a pancreatic beta cell and by providing a composition  
CC comprising an agent that reduces levels of fatty acids in the cells  
CC and protects beta-cells of the subject against lipid-mediated cell  
CC death. Cells can also be protected against nitric oxide mediated  
CC cytotoxicity by introducing into the cell an antioxidant agent.  
CC The methods can be used for protecting cells against immunotoxicity  
CC mediated by, e.g. IL-1 beta, IL-3, IL-5, IL-7, IL-9, IL-14, IL-17,  
CC gamma, IL-8, IL-2, IL-6, IL-2, IL-3, IL-5, IL-7, IL-9, IL-14, IL-17,  
CC granulocyte-macrophage colony stimulating factor or monocyte  
CC chemoattractant protein-1. The methods can be used for the treatment  
CC of e.g. insulin-dependent diabetes mellitus (IDDM), NIDDM, obesity,  
CC wasting syndromes, short stature, osteoporosis, inflammatory  
CC diseases, autoimmune diseases, or neurodegenerative diseases.  
SQ Sequence 1146 AA;

Query Match 71.4%; Score 45; DB 1; Length 1146;  
Best Local Similarity 85.7%; Pred. No. 8, 92e+01;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 80 SSPRHVR 86  
QY 120 SSPKHVR 126  
|||||

RESULT 6  
ID W36113 standard; Protein; 1153 AA.  
AC W36113, 1998 (first entry)  
DT Human inducible nitric oxide synthetase.  
DE Nitric oxide synthetase; NOS; iNOS; HPI NOS; human;  
KW erectile dysfunction; impotence; gene therapy; corpora cavernosa;  
KW relaxant.  
OS Homo sapiens.  
PN WO9742965-A1.  
PD 20-NOV-1997.  
PF 09-NOV-1997; U07643.  
PR 10-MAY-1996; US-01373.  
PA (GONZ/) GONZALEZ-CADAVID N F.  
PA (RAJF/) RAJFER J.  
PI Gonzalez-Cadavid NF, Rajfer J;  
DR WPI: 98-008577/01.  
DR N-PSDB: T98199.  
PT Treatment of erectile dysfunction - by introducing an agent into  
PT penile tissue, particularly for inducing cavernosal smooth muscle  
PT relaxation or increasing NOS levels  
PS Claim 12; Page 38-41, 53pp; English.  
CC This protein comprises human penis' inducible nitric oxide  
CC synthetase (HPI NOS). Its amino acid sequence was deduced from a  
CC cDNA clone (see T98199) derived from human penile smooth muscle  
CC cell mRNA. The invention is directed to a method of treating  
CC erectile dysfunction in a patient by providing an agent capable of  
CC treating erectile dysfunction, and introducing an effective amount  
CC of the agent into the penile tissue of the patient. Preferably,  
CC the agent induces cavernosal smooth muscle relaxation, and/or  
CC produces an increase in the level of NOS in tissue. Preferably,  
CC the NOS is iNOS, and the agent is introduced into the corpora  
CC cavernosa of the penis. The agent is preferably an NOS inducer, an  
CC NOS protein such as HPI NOS, a cDNA encoding an NOS such as HPI NOS,  
CC or cDNA transformed penile cells, especially corpora cavernosa  
CC cells.  
SQ Sequence 1153 AA;

Query Match 71.4%; Score 45; DB 1; Length 1153;  
Best Local Similarity 85.7%; Pred. No. 8, 92e+01;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 80 SSPRHVR 86  
QY 120 SSPKHVR 126  
|||||

RESULT 7  
ID R55764 standard; Protein; 1153 AA.  
AC R55764;  
DT 28-DEC-1994 (first entry)  
DE Sequence encoded by the cDNA clone for human hepatocyte inducible  
DE nitric oxide synthase.  
DE Nitric oxide synthase; hepatocyte; hypotensive shock; therapy.  
KW Homo sapiens.  
OS Homo sapiens.  
PN WO9412645-A.  
PD 09-JUN-1994.  
PF 23-NOV-1993; US-981344.  
PR 25-NOV-1992; US-981344.  
PA (UYPI-) UNIV PITTSBURGH.  
PI Billiar TR, Geller DA, Nussler AK, Simmons RL;  
DR WPI: 94-200273/24.  
DR N-PSDB: Q66914.  
PT cDNA clone encoding human inducible nitric oxide synthase - used  
PT to prevent the hypotensive shock seen with sepsis.  
PS Disclosure: Fig 1; 53pp; English.  
CC Q66914 is from human hepatocyte inducible nitric oxide synthase cDNA  
CC clone PHINOS from lambda Zap II cDNA library. The original source  
CC was induced human hepatocyte RNA. HINOS cDNA plasmid is pref.  
CC transformed in E. coli SOLR (ATCC 69126). The inventors claim a  
CC clone with the cDNA sequence in Q66914 and a cDNA clone which  
CC encodes R55764. The cloning and expression of a human tissue nitric  
CC oxide synthase cDNA provides a source of the enzyme for therapeutic  
CC purposes, for example to prevent the hypotensive shock seen with  
CC sepsis.  
SQ Sequence 1153 AA;

Query Match 71.4%; Score 45; DB 1; Length 1153;  
Best Local Similarity 85.7%; Pred. No. 8, 92e+01;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 80 SSPRHVR 86  
QY 120 SSPKHVR 126  
|||||

RESULT 8  
ID R88464 standard; Protein; 1153 AA.  
AC R88464;  
DT 13-MAY-1996 (first entry)  
DE Hepatocyte inducible nitric oxide synthase.  
DE Inducible nitric oxide synthase; iNOS; hepatocyte; gene therapy;  
KW vascular occlusive disease; cancer; infection.  
OS Homo sapiens.  
PN WO9600006-A1.  
PD 04-JAN-1996.  
PF 20-JUN-1995; U07849.  
PR 24-JUN-1994; US-265046.  
PA (UYPI-) UNIV PITTSBURGH.  
PI Billiar TR, Geller DA, Nussler AK, Simmons RL, Tzeng E;  
DR WPI: 96-068641/07.  
DR N-PSDB: T10115.  
PT Inducible nitric oxide synthase gene - useful in gene therapy to  
PT treat, e.g. vascular occlusive disease and cancer  
PS Disclosure: Page 54-58; 91pp; English.  
CC An inducible nitric oxide synthase (iNOS = R88464) is the product  
CC of a cDNA clone (T10115) derived from human hepatocytes induced  
CC for iNOS biosynthesis. The iNOS can be obt'd. by expression of  
CC the cDNA e.g. in mammalian host cells and is used in the  
CC development of selective inhibitors of NOS or to treat diseases

CC affected by nitric oxide.  
SQ Sequence 1153 AA;

Query Match 71.4%; Score 45; DB 1; Length 1153;  
Best Local Similarity 85.7%; Pred. No. 8.92e+01;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Db 80 SSPRHVR 86  
|||:||||  
QY 120 SSPKHVR 126

## RESULT 9

ID R63206 standard; Protein; 1153 AA.

AC R63206;  
DT 09-MAY-1995 (first entry)  
DE Nitric-oxide-synthase.  
KW Nitric-oxide-synthase; NOS; chondrocyte;  
KW interleukin-1-beta; PBHSINOS; arthritis; hypertension;  
KW septic shock; inflammation; ischemia; dementia; obesity; tumor;  
KW agonist; antagonist; vector; CHO; Chinese hamster ovary;  
KW cell culture.  
OS Homo sapiens.  
PN WO9423038-A.  
PD 13-OCT-1994.  
PF 25-MAR-1994; G00621.  
PA 26-MAR-1993; GB-006386.  
PI (WELL ) WELLCOME FOUND LTD.  
PI Charles IG, Moncada SE, Palmer RMJ, Moncada S;  
PI WPI: 94-333198/41.  
DR N-PSDB: Q77700.  
PT New human inducible nitric oxide synthase - useful for  
PT identifying enzyme inhibitors and stimulators, and for diagnosis  
PT and treatment of e.g. viral infections or tumours  
PS Disclosure: Page 25-31; 42pp; English.  
CC Human chondrocytes were incubated with interleukin-1-beta to induce  
CC nitric-oxide-synthase. cDNA was generated and used to construct a  
CC library in lambda ZAPII. This was screened with a 650 bp fragment of  
CC mouse inducible NO-synthase cDNA to identify the full-length clone  
CC PBHSINOS. The insert was transferred to pSVL to give a vector  
CC capable of expressing NO-synthase in CHO cells under control of a  
CC heterologous constitutive promoter.  
SQ Sequence 1153 AA;

Query Match 71.4%; Score 45; DB 1; Length 1153;  
Best Local Similarity 85.7%; Pred. No. 8.92e+01;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 80 SSPRHVR 86  
|||:||||  
QY 120 SSPKHVR 126

## RESULT 10

ID W38202 standard; Peptide; 43 AA.

AC W38202;  
DT 11-MAY-1998 (first entry)  
DE Arabidopsis SCARECROW SCR VHID domain.  
KW SCARECROW; SCR; transgenic plant; root; gravitropism;  
KW crop improvement.  
OS Arabidopsis thaliana.  
PN WO9741152-A1.  
PD 06-NOV-1997.  
PF 25-APR-1997; U07022.  
PA 26-APR-1996; US-842445.  
PI (CNY ) UNIV NEW YORK STATE.  
PI Benfey PN, Dilaurenzio L, Helariutta Y, Malamy JE,  
PI Pysh L, Wysocka-Diller J;  
PI WPI: 97-549683/50.  
PT DNA encoding Arabidopsis SCARECROW protein - useful to modify plant  
PT cell division and therefore alter root development, or alter plant  
PT stem or hypocotyl gravitropism

PS Disclosure: Fig 5D: 221pp; English.

CC This sequence comprises the VHID domain of Arabidopsis SCARECROW  
CC (SCR) protein (see W38178). SCR is a member of a novel protein  
CC family that contains a number of functional domains similar to  
CC those found in transcription factors. It is expressed specifically  
CC in embryo root progenitor tissue and in certain root and stem  
CC tissues. SCR regulates a specific asymmetric division, and  
CC controls gravitropic response in aerial structures and root  
CC formation. Translations of EST clones encoding putative peptides  
CC (see W38203-07) having similarities to the SCR VHID domain  
CC are provided. These are from Arabidopsis, maize and rice.  
CC Transgenic plants are claimed in which SCR expression is altered,  
CC thereby altering root development, or stem or hypocotyl  
CC gravitropism.  
SQ Sequence 43 AA;

Query Match 69.8%; Score 44; DB 1; Length 43;  
Best Local Similarity 55.6%; Pred. No. 1.18e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 31 RPPGPPHVR 39  
|||:||||  
QY 118 KPSPKHVR 126

## RESULT 11

ID Y02653 standard; Protein; 265 AA.

AC Y02653;  
DT 11-JUN-1999 (first entry)  
DE Human secreted protein encoded by gene 4 clone HCAA63.  
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
OS Homo sapiens.  
PN WO9902546-A1.  
PD 21-JAN-1999.

PF 07-JUL-1998; U13684.  
PR 12-SEP-1997; US-058785.  
PR 08-JUL-1997; US-051916.  
PR 08-JUL-1997; US-051918.  
PR 08-JUL-1997; US-051919.  
PR 08-JUL-1997; US-051920.  
PR 08-JUL-1997; US-051925.  
PR 08-JUL-1997; US-051926.  
PR 08-JUL-1997; US-051928.  
PR 08-JUL-1997; US-051929.  
PR 08-JUL-1997; US-051930.  
PR 08-JUL-1997; US-051931.  
PR 08-JUL-1997; US-051932.  
PR 08-JUL-1997; US-052732.  
PR 08-JUL-1997; US-052733.  
PR 08-JUL-1997; US-052793.  
PR 08-JUL-1997; US-052795.  
PR 08-JUL-1997; US-052803.  
PR 18-AUG-1997; US-055684.  
PR 18-AUG-1997; US-055722.  
PR 18-AUG-1997; US-055723.  
PR 18-AUG-1997; US-055947.  
PR 18-AUG-1997; US-055948.  
PR 18-AUG-1997; US-055949.  
PR 18-AUG-1997; US-055950.  
PR 18-AUG-1997; US-055953.  
PR 18-AUG-1997; US-055954.  
PR 18-AUG-1997; US-055964.  
PR 18-AUG-1997; US-055984.  
PR 18-AUG-1997; US-056360.  
PR 12-SEP-1997; US-058660.  
PR 12-SEP-1997; US-058661.



12-SEP-1997; US-058664.  
 (HUMA-) HUMAN GENOME SCI INC.  
 PA Brewer LA, Ener R, Fischer CL, Kyaw H, Lafleur DW, Li Y, Moore PA,  
 PI Olsen HS, Rosen CA, Ruben SM, Shi Y, Soppet DR, Zeng Z;  
 PI WPI: 99-120770/10.  
 DR N-PSDB: X27314.  
 PT New isolated human genes and the secreted polypeptides they encode -  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders  
 PS Claim 11: Page 320-321: 464pp; English.  
 CC This sequence represents a secreted human protein encoded by the gene  
 CC clone detailed in the descriptor line. The gene can be used to generate  
 CC fusion proteins by linking to the gene to a human immunoglobulin FC  
 CC portion (e.g. X27302) for increasing the stability of the fused protein  
 CC as compared to the human protein only.  
 CC The invention relates to 123 novel genes and their fragments (nucleic  
 CC acid sequences: X27311-X27449; amino acid sequences Y02650-Y02788) which  
 CC are useful for preventing, treating or ameliorating medical conditions  
 CC e.g. blood protein or gene therapy. Also, pathological conditions can be  
 CC diagnosed by determining the amount of the new polypeptides in a sample  
 CC or by determining the presence of mutations in the new polynucleotides.  
 CC Specific uses are described for each of the 123 polynucleotides, based on  
 CC which tissues they are most highly expressed in (see X27311 for described  
 CC uses).  
 CC Sequence 265 AA;  
 SQ

Query Match 69.8%; Score 44; DB 1; Length 265;  
 Best Local Similarity 62.5%; Pred. No. 1.18e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Db 7 PPSRHLR 14  
 QY 119 PSSPKHVR 126  
 I:|||||  
 I:|||||

RESULT 12 standard; Protein: 350 AA.  
 ID W80397; 1999 (first entry)  
 AC W80397; 1999 (first entry)  
 DE A secreted protein; immune stimulating; suppressing;  
 KW Haematopoiesis regulating activity; tissue growth; activity; activin;  
 KW Inhibin activity; chemotactic; chemokinetic activity; haemostatic;  
 KW Thrombolytic activity; anti-inflammatory activity; cadherin;  
 KW Tumour invasion suppressor activity; tumour inhibition activity.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Misc\_difference 286 /note= "not defined"  
 FT W09840486-A2.  
 PN W09840486-A2.  
 PD 17-SEP-1998.  
 PF 13-MAR-1998; 004977.  
 PR 29-OCT-1997; US-960022.  
 PR 14-MAR-1997; US-815047.  
 PA (GEM ) GENETICS INST INC.  
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,  
 PI Racie LA, Spaulding V, Treacy M;  
 PI WPI: 98-520812/44.  
 DR N-PSDB: V63171.  
 PT New isolated human polynucleotide(s) and secreted proteins -  
 PT obtained from e.g. human foetal kidney, placenta, foetal brain,  
 PT adult testes, adult brain or adult uterus cDNA libraries  
 PS Claim 20: Pages 71-72: 110pp; English.  
 CC The present sequence represents a secreted protein. The nucleic acid  
 CC probe is isolated from a human adult placenta cDNA library using  
 CC probe V63181. The polypeptide may have biological activities such as  
 CC e.g. nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity or other activities.  
 CC Sequence 350 AA;  
 SQ

Query Match 69.8%; Score 44; DB 1; Length 350;  
 Best Local Similarity 75.0%; Pred. No. 1.18e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Db 57 KNSPKHV 64  
 QY 118 KPSSPKHV 125  
 I:|||||  
 I:|||||

RESULT 13  
 ID W38178 standard; Protein: 653 AA.  
 AC W38178;  
 DT 11-MAY-1998 (first entry)  
 DE Arabidopsis SCARECROW protein.  
 KW SCARECROW; SCR gene; transgenic plant; root; gravitropism;  
 KW crop improvement.  
 OS Arabidopsis thaliana.  
 FH Key Location/Qualifiers  
 FT Region 1..267  
 FT /note= "N-terminal region rich in Gln, Ser, Thr and  
 FT Pro homopolymetric structures"  
 FT Peptide 291..322  
 FT /note= "leucine heptad repeat"  
 FT Peptide 436..473  
 FT /note= "leucine heptad repeat"  
 FT Domain 58..99  
 FT /label= MOTIF-I  
 FT Domain 265..341  
 FT /label= MOTIF-II  
 FT Domain 373..435  
 FT /label= MOTIF-III(VHIID)  
 FT Domain 436..480  
 FT /label= MOTIF-IV  
 FT Domain 457..496  
 FT /label= MOTIF-V  
 FT Domain 516..649  
 FT /label= MOTIF-VI  
 PN W09741152-A1.  
 PD 06-NOV-1997.  
 PF 25-APR-1997; U07022.  
 PR 24-APR-1997; US-842445.  
 PR 26-APR-1996; US-638617.  
 PA (UYN ) UNIV NEW YORK STATE.  
 PI Benfey PN, Dilaurenzio L, Helariutta Y, Malamy JE,  
 PI Pysh L, Wysocka-Diller J;  
 PI WPI: 97-549683/50.  
 DR N-PSDB: T95753.  
 PT DNA encoding Arabidopsis SCARECROW protein - useful to modify plant  
 PT cell division and therefore alter root development, or alter plant  
 PT stem or hypocotyl gravitropism.  
 PS Claim 12: Page 92-94: 221pp; English.  
 CC This is the deduced amino acid sequence of the Arabidopsis  
 CC SCARECROW (SCR) gene (see T95753). SCR is a member of a novel  
 CC protein family and contains a number of potential functional  
 CC domains similar to those found in transcription factors. SCR is  
 CC expressed specifically in embryo root progenitor tissue and in  
 CC certain root and stem tissues. It regulates a series of asymmetric  
 CC division, and controls gravitropic response in aerial structures  
 CC and root formation. SCR proteins (see W38178-201) from dicots and  
 CC monocots. SCR proteins lacking 1-4 of MOTIFS I-VI, and polypeptides  
 CC corresponding to MOTIF I, II, III, IV, V or VI of SCR are claimed.  
 CC so that cell division can be engineered to overexpress the SCR protein,  
 CC transgenic plants can be engineered to increase root growth, resulting in thicker  
 CC root development, while a plant with an altered stem or hypocotyl  
 CC gravitropism is less susceptible to lodging. Plants that contain  
 CC an antisense molecule that suppresses the expression of endogenous  
 CC SCR gene product have thinner root development. A gene of interest  
 CC can be placed under control of a SCR promoter and expressed in a  
 CC plant to confer herbicide, salt, pathogen or insect resistance, or  
 CC when expressed in stems to increase starch, lignin or cellulose  
 CC biosynthesis (all claimed).  
 CC Sequence 653 AA;  
 SQ

Query Match 69.8%; Score 44; DB 1; Length 653;  
 Best Local Similarity 55.6%; Pred. No. 1.18e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 424 RGGPPHVR 432  
 QY 118 KPSSPKHVR 126

## RESULT 14

ID W20051 standard; Protein; 657 AA.  
 AC W20051;  
 DT 10-SEP-1997 (first entry)  
 DE Methionyl-tRNA synthetase from Staph. aureus.  
 KW RNA synthetase; Bacillus subtilis; metS; immunological response;  
 KW antibody; bacterial infection; adherence; damaged tissue;  
 KW wound healing; skin; protection; vaccine.  
 OS Staphylococcus aureus.  
 PN EP-785288-Al.  
 PD 23-JUL-1997.  
 PF 17-JAN-1997; 300317.  
 PR 19-JAN-1996; GB-001095.  
 PR 27-JUL-1996; GB-015845.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 PI Hodgson JE, Lawlor EJ;  
 DR WPI; 97-365943/34.  
 DR N-PSDB; T71309.  
 PT DNA encoding methionyl-tRNA synthetase from Staphylococcus aureus  
 PS WCUH 29 - useful for protection against bacterial infections  
 CC Claim 1: Pages 19-20; 31pp; English.  
 CC The present sequence is a novel methionyl tRNA synthetase  
 CC protein from Staphylococcus aureus strain WCUH29 (NCIMB 40771),  
 CC which is related by amino acid sequence homology to Bacillus  
 CC subtilis tRNA synthetase encoded by the metS gene. The enzyme  
 CC catalyses the aminoacylation of tRNA-Met, by a two step mechanism.  
 CC The first step involves formation of a stable enzyme:methionyl  
 CC adenylylate complex resulting from the specific binding and reaction  
 CC of ATP and L-methionine. Subsequently, the 3' terminal adenosine of  
 CC enzyme-bound tRNA-Met reacts with the aminoacyl adenylylate, leading  
 CC to the esterification of the tRNA and release of AMP. Vectors  
 CC comprising the DNA (or polynucleotides having at least 70 %  
 CC identity to it) can be used for the recombinant production of the  
 CC enzyme. The enzyme or its related DNA (through gene therapy) is  
 CC used to induce an immunological response in a mammal to generate  
 CC antibodies to protect against disease. The antibodies protect  
 CC against invasion of bacteria, e.g. by blocking adherence of  
 CC bacteria to damaged tissue, including wounds in skin or connective  
 CC tissue caused by mechanical, chemical or thermal damage or by  
 CC implantation of in-dwelling devices, or wounds in the mucous  
 CC membranes.  
 SQ Sequence 657 AA;

Query Match 69.8%; Score 44; DB 1; Length 657;  
 Best Local Similarity 85.7%; Pred. No. 1.18e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 227 PSNPKHV 233  
 QY 119 PSSPKHV 125

## RESULT 15

ID W27660 standard; Protein; 665 AA.  
 AC W27660;  
 DT 28-APR-1998 (first entry)  
 DE Streptococcus pneumoniae methionyl tRNA synthetase.  
 KW Methionyl tRNA synthetase; metS polypeptide; antibacterial; vaccine;  
 KW immune response; meningitis; bacterial infection.  
 OS Streptococcus pneumoniae.  
 PN W09739012-Al.  
 PD 23-OCT-1997.  
 PF 18-APR-1997; U06555.

PR 18-APR-1996; GB-007999.  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PI (SMIK ) SMITHKLINE BEECHAM PLC.  
 PI Lawlor EJ;  
 DR WPI; 97-526388/48.  
 DR N-PSDB; T88098.  
 PT Polynucleotide encoding Streptococcus pneumoniae methionyl tRNA  
 PT synthetase - useful to diagnose, treat or prevent bacterial  
 PT infection, especially meningitis  
 PS Claim 12; Page 36-38; 47pp; English.  
 CC The present sequence represents methionyl tRNA synthetase (metS) from  
 CC Streptococcus pneumoniae. The polynucleotide encoding metS can be used  
 CC for the recombinant production of metS, which can be used to treat  
 CC conditions requiring metS activity. Sequences antisense to the  
 CC polynucleotide can be used to control metS expression. MetS, or metS  
 CC expressing vectors can be used to induce an immune response, i.e. an  
 CC antibody (Ab) and/or T cell response, against S. pneumoniae to protect  
 CC against infection, or to screen for antagonists or agonists of the  
 CC polynucleotide encoding metS or metS's activity, i.e. antibacterials.  
 CC The antagonist, e.g. an anti-metS Ab, can be used to treat conditions  
 CC requiring metS inhibition, e.g. an S. pneumoniae infection, particularly  
 CC meningitis. A fragment of the polynucleotide encoding metS can be used as a  
 CC probe to isolate full length or related sequences, or diagnose, e.g. by  
 CC polymerase chain reaction, infection stage and type, including mutation  
 CC and polymorphism detection. Diagnosis may also be achieved by detecting  
 CC metS gene overexpression, e.g. by immunoassay. The Ab can be used to treat  
 CC infection, isolate or identify metS expressing clones, purify metS and  
 CC as an immunoassay reagent. More generally, the products can prevent  
 CC adhesion of bacteria to wounds and in dwelling devices, block metS  
 CC protein mediated invasion of mammalian cells and block the normal  
 CC progression of infection.  
 SQ Sequence 665 AA;

Query Match 69.8%; Score 44; DB 1; Length 665;  
 Best Local Similarity 85.7%; Pred. No. 1.18e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 227 PSNPKHV 233  
 QY 119 PSSPKHV 125

Search completed: Wed May 10 13:07:15 2000  
 Job time : 8 secs.

\*\*\*\*\*  
WPSRELH  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 13:07:33 2000; MasPar time 45.55 Seconds  
Tabular output not generated. 2.560 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (118-126) from US09376430A.pap (15 of 25)  
Perfect Score: 63  
Sequence: 1 KPSSPKHVR 9

Scoring table: PAM 150  
Gap 11  
Searched: 131253 seqs, 12956647 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-issued  
1:5A\_COMB 2:5B\_COMB 3:PCT\_COMB 4:backfiles1  
Statistics: Mean 14.161; Variance 39.985; scale 0.354  
pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	47	74.6	1324	2	US-08-811- Sequence 56, Applicatio	3.20e+01
2	45	71.4	1153	3	PCT-US93-1 Sequence 2, Applicatio	5.52e+01
3	45	71.4	1153	2	US-08-465- Sequence 2, Applicatio	5.52e+01
4	45	71.4	1153	2	US-08-465- Sequence 2, Applicatio	5.52e+01
5	45	71.4	1153	3	PCT-US93-0 Sequence 2, Applicatio	5.52e+01
6	45	71.4	1153	3	US-08-314- Sequence 2, Applicatio	5.52e+01
7	44	69.8	350	2	US-08-960- Sequence 8, Applicatio	7.24e+01
8	44	69.8	895	1	US-08-123- Sequence 8, Applicatio	7.24e+01
9	44	69.8	895	1	US-08-483- Sequence 8, Applicatio	7.24e+01
10	42	66.7	574	2	US-08-906- Sequence 2, Applicatio	1.24e+02
11	42	66.7	992	1	US-08-482- Sequence 1, Applicatio	1.24e+02
12	42	66.7	992	1	US-08-127- Sequence 1, Applicatio	1.24e+02
13	42	66.7	1135	2	US-08-469- Sequence 97, Applicatio	1.24e+02
14	41	65.1	588	2	US-08-381- Sequence 6, Applicatio	1.61e+02
15	41	65.1	588	3	PCT-US93-0 Sequence 12, Applicatio	1.61e+02
16	41	65.1	588	2	US-08-469- Sequence 6, Applicatio	2.10e+02
17	40	63.5	800	2	US-08-691- Sequence 78, Applicatio	2.10e+02
18	40	63.5	800	2	US-08-469- Sequence 72, Applicatio	2.10e+02
19	40	63.5	905	2	US-08-574- Sequence 9, Applicatio	2.10e+02
20	40	63.5	937	2	US-08-469- Sequence 105, Applicat	2.10e+02
21	40	63.5	1135	2	US-08-574- Sequence 7, Applicatio	2.10e+02
22	40	63.5	1135	2	US-07-852- Sequence 2, Applicatio	2.10e+02
23	40	63.5	2206	1		

RESULT ID	US-08-811-897A-56	STANDARD	PRT	1324 AA.
XX	xxxxxx			
AC				
XX				
DT				
XX				
DE				
XX				
CC	Sequence 56, Application US/08811897A			
CC	Sequence 56, Application US/08811897A			
CC	Patent No. 5858787			
CC	GENERAL INFORMATION:			
CC	APPLICANT: ONDA, Haruo			
CC	APPLICANT: OHTAKI, Tetsuya			
CC	APPLICANT: MASUDA, Yasushi			
CC	APPLICANT: KITADA, Chieko			
CC	APPLICANT: ISHIBASHI, Yoshihiro			
CC	APPLICANT: HOSoya, Masaki			
CC	APPLICANT: OGI, Kazuhiro			
CC	APPLICANT: MIYAMOTO, Yasunori			
CC	APPLICANT: HABATA, Yugo			
CC	APPLICANT: SHIMAMOTO, No. 5858787			
CC	TITLE OF INVENTION: PACAP RECEPTOR PROTEIN, METHOD FOR PREPARING			
CC	TITLE OF INVENTION: SAID PROTEIN, AND USE THEREOF			
CC	NUMBER OF SEQUENCES: 56			
CC	CORRESPONDENCE ADDRESS:			
CC	ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &			
CC	ADDRESSER: CUSHMAN			
CC	STREET: 130 Water Street			
CC	CITY: Boston			
CC	STATE: Massachusetts			
CC	COUNTRY: US			
CC	ZIP: 02109			
CC	COMPUTER READABLE FORM:			
CC	MEDIUM TYPE: Diskette			
CC	COMPUTER: IBM Compatible			
CC	OPERATING SYSTEM: DOS			
CC	SOFTWARE: FASTSEQ for Windows Version 2.0			
CC	CURRENT APPLICATION DATA:			
CC	APPLICATION NUMBER: US/08/811,897A			
CC	FILING DATE: 05-MAR-1997			
CC	CLASSIFICATION: 435			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: 08/202,986			
CC	FILING DATE: February 25, 1994			
CC	ATTORNEY/AGENT INFORMATION:			
CC	NAME: RESNICK, David S.			

CC REGISTRATION NUMBER: 34235  
CC REFERENCE/DOCKET NUMBER: 44168-DIV  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (617)523-3400  
CC TELEFAX: (617)523-6440 56:  
CC INFORMATION FOR SEQ ID NO:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1324 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC SEQUENCE 1324 AA; 149660 MW; 9039953 CN;  
  
Query Match 74.4%; Score 47; DB 2; Length 1324;  
Best Local Similarity 66.7%; Pred. No. 3.20e+01;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
Db 4 RPPSPHVR 12  
QY 118 KPSPKHVR 126  
:|||  
  
RESULT 2  
ID PCT-US93-11401-2 STANDARD; PRT: 1153 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE Sequence 2, Application PCTUS9311401  
XX  
CC Sequence 2, Application PCTUS9311401  
CC GENERAL INFORMATION:  
CC APPLICANT: Billiar, Timothy R.  
CC APPLICANT: Nussier, Andreas K.  
CC APPLICANT: Geller, David A.  
CC APPLICANT: Simmons, Richard L.  
CC TITLE OF INVENTION: cDNA clone for Human Inducible Nitric  
CC OXIDE SYNTHASE AND PROCESS FOR PREPARING SAME  
CC NUMBER OF SEQUENCES: 2  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Arnold B. Silverman  
CC ADDRESSEE: Eckert Seamas Cherin & Mellott  
CC STREET: 600 Grant Street, 42nd Floor  
CC CITY: Pittsburgh  
CC STATE: PA  
CC COUNTRY: USA  
CC ZIP: 15219  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US93/11401  
CC FILING DATE: 25-NOV-1992  
CC CLASSIFICATION:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US/07/991,344  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Silverman, Arnold B.  
CC REGISTRATION NUMBER: 22,614  
CC REFERENCE/DOCKET NUMBER: 116972  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (412) 566-6000  
CC TELEFAX: (412) 566-6099  
CC TELEX: 866172  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1153 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear

CC MOLECULE TYPE: protein  
SQ SEQUENCE 1153 AA; 131027 MW; 7024056 CN;  
  
Query Match 71.4%; Score 45; DB 3; Length 1153;  
Best Local Similarity 85.7%; Pred. No. 5.52e+01;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Db 80 SSPRHVR 86  
QY 120 SSPKHVR 126  
|||  
  
RESULT 3  
ID US-08-465-522-2 STANDARD; PRT: 1153 AA.  
XX  
AC xxxxxx  
XX  
DT  
XX  
DE Sequence 2, Application US/08465522  
XX  
CC Sequence 2, Application US/08465522  
CC Patent No. 5882908  
CC GENERAL INFORMATION:  
CC APPLICANT: Billiar, Timothy R.  
CC APPLICANT: Nussier, Andreas K.  
CC APPLICANT: Geller, David A.  
CC APPLICANT: Simmons, Richard L.  
CC TITLE OF INVENTION: cDNA Clone for Human Inducible Nitric  
CC OXIDE SYNTHASE AND PROCESS FOR PREPARING SAME  
CC NUMBER OF SEQUENCES: 2  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Lewis F. Gould, Jr.  
CC ADDRESSEE: Eckert Seamas Cherin & Mellott  
CC STREET: 1700 Market St. Suite 3232  
CC CITY: Philadelphia  
CC STATE: PA  
CC COUNTRY: USA  
CC ZIP: 19103  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/465,522  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Gould, Jr., Lewis F.  
CC REGISTRATION NUMBER: 25,057  
CC REFERENCE/DOCKET NUMBER: 116972-6  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (215) 575-6020  
CC TELEFAX: (215) 575-6015  
CC TELEX:  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1153 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 1153 AA; 131027 MW; 7024056 CN;  
  
Query Match 71.4%; Score 45; DB 2; Length 1153;  
Best Local Similarity 85.7%; Pred. No. 5.52e+01;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Db 80 SSPRHVR 86  
QY 120 SSPKHVR 126  
|||

Thu May 11 06:49:58 2000

US-09-376-430-2-15-rai

RESULT 4 US-08-265-046-2 STANDARD; PRT; 1153 AA.

XXXXXX

Sequence 2, Application US/08265046

Sequence 2, Application US/08265046

Patent No. 5658565

GENERAL INFORMATION:

APPLICANT: Timothy R. Balliar

APPLICANT: Edith Tzeong Nussler

APPLICANT: Andreas K. Geller

APPLICANT: David A. Geller

APPLICANT: Richard L. Simmons

TITLE OF INVENTION: Inducible Nitric Oxide Synthase

TITLE OF INVENTION: Gene for Treatment of Disease

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lewis F. Gould, Jr.

ADDRESSEE: Eckert Seamans Cherin & Mellott

STREET: 1700 Market Street, Suite 3232

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/265,046

FILING DATE: 24-JUN-1994

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Gould, Lewis F. Jr.

REGISTRATION NUMBER: 25,057

REFERENCE/DOCKET NUMBER: 119130

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 575-6020

TELEFAX: (215) 575-6015

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1153 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 1153 AA; 131027 MW; 7024056 CN;

Query Match 71.4%; Score 45; DB 1; Length 1153;

Best Local Similarity 85.7%; Pred. No. 5.52e+01;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 80 SSPKVR 86

111:111

QY 120 SSPKVR 126

RESULT 5

PC-TUS95-07849-2 STANDARD; PRT; 1153 AA.

XXXXXX

Sequence 2, Application PC/TUS9507849

Sequence 2, Application PC/TUS9507849

GENERAL INFORMATION:

CC APPLICANT: University of Pitt

CC APPLICANT: Education

CC TITLE OF INVENTION: Inducible N

CC TITLE OF INVENTION: Gene for Tre

CC NUMBER OF SEQUENCES: 2

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Lewis F. Gould, Jr.

CC ADDRESSEE: Eckert Seamans Cherin & M

CC STREET: 1700 Market Street, Suite 3232

CC CITY: Philadelphia

CC STATE: PA

CC COUNTRY: USA

CC ZIP: 19103

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US95/07849

CC FILING DATE:

CC CLASSIFICATION:

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Gould, Lewis F. Jr.

CC REGISTRATION NUMBER: 25,057

CC REFERENCE/DOCKET NUMBER: 119130-2

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (215) 575-6020

CC TELEFAX: (215) 575-6015

CC TELEX:

CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 1153 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

CC SEQUENCE 1153 AA; 131027 MW; 7024056 CN;

Query Match 71.4%; Score 45; DB 3; Length 1153;

Best Local Similarity 85.7%; Pred. No. 5.52e+01;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 80 SSPKVR 86

111:111

QY 120 SSPKVR 126

RESULT 6

US-08-314-917-2 STANDARD; PRT; 1153 AA.

XXXXXX

Sequence 2, Application US/08314917

Sequence 2, Application US/08314917

Patent No. 5468630

GENERAL INFORMATION:

APPLICANT: Billiar, Timothy R.

APPLICANT: Nussler, Andreas K.

APPLICANT: Geller, David A.

APPLICANT: Simmons, Richard L.

TITLE OF INVENTION: cDNA Clone for Human Inducible Nitric

TITLE OF INVENTION: Oxide Synthase And Process for Preparing Same

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold B. Silverman

ADDRESSEE: Eckert Seamans Cherin & Mellott

STREET: 600 Grant

CITY: Pittsburgh

STATE: PA

COUNTRY: USA

CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/960,022  
 CC FILING DATE:  
 CC CLASSIFICATION: 514  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Sprunger, Suzanne A.  
 CC REGISTRATION NUMBER: 41,323  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (617) 498-8284  
 CC TELEFAX: (617) 876-5851  
 CC INFORMATION FOR SEQ ID NO: 8:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 350 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS:  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 350 AA; 39042 MW; 660281 CN;  
 CC  
 CC Query Match 69.8%; Score 44; DB 2; Length 350;  
 CC Best Local Similarity 75.0%; Pred. No. 7.24e-01;  
 CC Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 Db 57 KNSGPKHV 64  
 QY 118 KPSSPKHV 125  
 CC  
 CC RESULT 8  
 CC ID US-08-123-161A-8 STANDARD: PRT: 895 AA.  
 CC XX  
 CC AC xxxxxx  
 CC XX  
 CC DT  
 CC XX  
 CC DE  
 CC XX  
 CC Sequence 8, Application US/08123161A  
 CC Sequence 8, Application US/0812  
 CC Patent No. 5449616  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Campbell, V  
 CC APPLICANT: Roberts,  
 CC APPLICANT: Andersc  
 CC APPLICANT: Ibrag  
 CC APPLICANT: Yag  
 CC TITLE OF INVENT  
 CC TITLE OF INV  
 CC NUMBER OF  
 CC CORRESPOND  
 CC ADDRESS  
 CC STREET:  
 CC CITY: YC  
 CC STATE: N  
 CC COUNTRY:  
 CC ZIP: 039  
 CC COMPUTER RE  
 CC MEDIUM TV  
 CC COMPUTER:  
 CC OPERATING  
 CC SOFTWARE:  
 CC CURRENT APPL  
 CC APPLICATION  
 CC FILING DAT  
 CC CLASSIFICA  
 CC PRIOR APPLIC  
 CC APPLICATION  
 CC FILING DAT  
 CC ATTORNEY/AGEN  
 CC NAME: Fair  
 CC REGISTRATIO  
 CC REFERENCE/DI  
 CC TELECOMMUNICA  
 CC TELEPHONE:

US-00-376-430-2-15.ra1

Thu May 11 06:49:58 2000

CC TELEFAX: (207) 363-0528  
 CC INFORMATION FOR SEQ ID NO: 8:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 895 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 895 AA; 97580 MW; 4414069 CN;

Query Match 69.8%; Score 44; DB 1; Length 895;  
 Best Local Similarity 66.7%; Pred. No. 7.24e-01;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 303 KPPLPKRVR 311  
 QY 118 KPSSPKHVR 126

RESULT 9  
 ID US-08-483-278-8 STANDARD; PRT: 895 AA.  
 XX  
 AC xxxxxx  
 XX  
 DT

Sequence 8, Application US/08483278

Sequence 8, Application US/08483278  
 Patent No. 5686073  
 GENERAL INFORMATION:  
 APPLICANT: Campbell, Kevin P.  
 APPLICANT: Ibraghimov, Oxana B.  
 APPLICANT: Ervasti, James M.  
 APPLICANT: Leveille, Cynthia J.  
 TITLE OF INVENTION: NUCLEIC ACID ENCODING DYSTROPHIN-ASSOCIATED  
 TITLE OF INVENTION: PROTEIN  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Kevin M. Farrell, P.C.  
 STREET: P.O. Box 999  
 CITY: York Harbor  
 STATE: ME  
 COUNTRY: USA  
 ZIP: 03911

COMPUTER READABLE FORM: disk  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/483,278  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/123,161  
 FILING DATE: 16-SEP-93  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Farrell, Kevin M.  
 REGISTRATION NUMBER: 35,505  
 REFERENCE/DOCKET NUMBER: UIRF89-11A5  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (207) 363-0558  
 TELEFAX: (207) 363-0528  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 895 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE 895 AA; 97580 MW; 4414069 CN;

Query Match 69.8%; Score 44; DB 1; Length 895;  
 Best Local Similarity 66.7%; Pred. No. 7.24e-01;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 303 KPPLPKRVR 311  
 QY 118 KPSSPKHVR 126

RESULT 10  
 ID US-08-906-713-2 STANDARD; PRT: 574 AA.  
 XX  
 AC xxxxxx  
 XX  
 DT

Sequence 2, Application US/08906713  
 Sequence 2, Application US/08906713  
 Patent No. 5965704  
 GENERAL INFORMATION:  
 APPLICANT: Lok, Si  
 APPLICANT: Adams, Robyn L.  
 APPLICANT: Jelmsberg, Anna C.  
 APPLICANT: Whitmore, Theodore E.  
 APPLICANT: Fairah, Theresa M.  
 TITLE OF INVENTION: MAMMALIAN ZCYTOR11  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Zymogenetics  
 STREET: 1201 Eastlake Ave East  
 CITY: Seattle  
 STATE: WA  
 COUNTRY: USA  
 ZIP: 98102

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/906,713  
 FILING DATE:  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lunn, Paul G  
 REGISTRATION NUMBER: 32,743  
 REFERENCE/DOCKET NUMBER: 97-52  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 206-442-6627  
 TELEFAX: 206-442-6678  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 574 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 FRAGMENT TYPE: internal  
 SEQUENCE 574 AA; 62977 MW; 1849163 CN;

Query Match 66.7%; Score 42; DB 2; Length 574;  
 Best Local Similarity 85.7%; Pred. No. 1.24e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Db 420 SSPKHLR 426  
 QY 120 SSPKHLR 126

RESULT 11







CC FILING DATE: 31-JUL-1992  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: FITCHER, EDMUND R  
 CC REGISTRATION NUMBER: 27,829  
 CC REFERENCE/DOCKET NUMBER: CRP-081CP  
 CC INFORMATION FOR SEQ ID NO: 12:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 588 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC FEATURE:  
 CC NAME/KEY: Protein  
 CC LOCATION: 1..588  
 CC OTHER INFORMATION: /note= "PRE-PRO-DPP"  
 CC PUBLICATION INFORMATION:  
 CC AUTHORS: PADGETT,  
 CC JOURNAL: NATURE  
 CC VOLUME: 325  
 CC PAGES: 81-84  
 CC DATE: 1987  
 CC  
 CC SEQUENCE 588 AA; 65833 MW; 1765810 CN;  
 CC  
 CC Query Match 65.1%; Score 41; DB 2; Length 588;  
 CC Best Local Similarity 55.6%; Pred. No. 1.61e+02;  
 CC Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 CC  
 CC Db 414 KPAPHHHVR 422  
 CC Qy 118 KPSSPKHVR 126

Search completed: Wed May 10 13:08:29 2000  
 Job time : 56 secs.

\*\*\*\*\*  
 W P S R L H  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 13:06:42 2000; Maspar time 3.98 Seconds  
 Tabular output not generated. 106.781 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (118-126) from US09376430A.pap (15 of 25)  
 Perfect Score: 63  
 Sequence: 1 KPSSPKHVR 9

Scoring table: PAM 150  
 Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: pir62  
 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 21.022; Variance 23.698; scale 0.887

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	52	82.5	6839	2	twitchin - Caenorhabd	2.11e-01
2	47	74.6	66	2	microtubule-associate	3.18e+00
3	47	74.6	205	2	proline-rich sheath p	3.18e+00
4	47	74.6	205	2	microfilament sheath p	3.18e+00
5	47	74.6	459	2	vasoactive intestinal	3.18e+00
6	46	73.0	120	2	ribosomal protein L22	5.36e+00
7	46	73.0	350	2	microtubule-associate	5.36e+00
8	46	73.0	1861	2	microtubule-associate	5.36e+00
9	45	71.4	144	2	hypothetical protein	8.96e+00
10	45	71.4	378	2	ferredoxin-NADP+ red	8.96e+00
11	45	71.4	528	2	[RNA-polymerase]-subu	8.96e+00
12	45	71.4	1153	2	nitric-oxide synthase	8.96e+00
13	44	69.8	296	2	hypothetical protein	1.49e+01
14	44	69.8	327	2	ferredoxin-NADP+ red	1.49e+01
15	44	69.8	332	2	regulatory protein -	1.49e+01
16	44	69.8	533	2	tyrosine phosphoprote	1.49e+01
17	44	69.8	576	2	calcium-dependent pro	1.49e+01
18	44	69.8	590	2	ovarian abundant mess	1.49e+01
19	44	69.8	895	2	dystroglycan - human	1.49e+01
20	44	69.8	1165	1	pol polyprotein - gib	2.45e+01
21	43	68.3	263	2	developmentally regul	2.45e+01
22	43	68.3	378	1	probable ferredoxin--	2.45e+01
23	43	68.3	536	2	glutamate--trna ligas	2.45e+01

24	43	68.3	610	1	A49082	calcium-dependent pro	2.45e+01
25	43	68.3	619	2	S54636	probable membrane pro	2.45e+01
26	43	68.3	691	2	T15983	hypothetical protein	2.45e+01
27	43	68.3	692	1	S57592	probable phosphoester	2.45e+01
28	43	68.3	932	2	S09151	suvar(317 protein - f	2.45e+01
29	43	68.3	1534	2	S59604	DNA (cytosine-5--)-met	2.45e+01
30	42	66.7	127	2	S37911	hypothetical protein	3.99e+01
31	42	66.7	288	2	A71148	hypothetical protein	3.99e+01
32	42	66.7	378	2	T02215	ferredoxin-NADP+ red	3.99e+01
33	42	66.7	378	2	T02977	ferredoxin-NADP+ red	3.99e+01
34	42	66.7	410	1	IXBEL0	alpha trans-inducing	3.99e+01
35	42	66.7	442	2	D71481	probable trigger fact	3.99e+01
36	42	66.7	494	2	G69433	signal-transducing hi	3.99e+01
37	42	66.7	541	1	MMBYH2	glucose transport pro	3.99e+01
38	42	66.7	560	2	S46724	hexose transport prot	3.99e+01
39	42	66.7	566	2	S51081	glucose transport pro	3.99e+01
40	42	66.7	852	2	T10811	channel associated pr	3.99e+01
41	42	66.7	870	2	G01974	channel associated pr	3.99e+01
42	42	66.7	895	2	S20582	dystrophin-associated	3.99e+01
43	42	66.7	992	1	GNWVR3	structural polyprotei	3.99e+01
44	42	66.7	1055	2	H70951	hypothetical protein	3.99e+01
45	42	66.7	1083	2	S54293	regulator protein pl2	3.99e+01

## ALIGNMENTS

RESULT 1  
 ENTRY S57242 #type complete  
 TITLE twitchin - Caenorhabditis elegans  
 ALTERNATE\_NAMES myosin-regulating protein  
 CONTAINS Protein kinase (EC 2.7.11-)  
 ORGANISM #formal\_name Caenorhabditis elegans  
 DATE 28-Oct-1995 #sequence\_revision 24-Oct-1997 #text\_change 18-Jun-1999  
 ACCESSIONS S57242; S07571; S06797; S57218  
 REFERENCE S57242  
 #authors Benian, G.M.; L'Hernault, S.W.; Morris, M.E.  
 #submission Submitted to the EMBL Data Library, February 1993  
 #description Additional sequence complexity within twitching of Caenorhabditis elegans muscle.  
 #accession S57242  
 #molecule\_type DNA  
 #residues 1-6839 ##label BEN1  
 ##cross-references EMBL:L10351  
 ##experimental\_source var. Bristol  
 #authors S07571  
 #submission Submitted to the EMBL Data Library, November 1989  
 #accession S07571  
 #molecule\_type DNA  
 #residues 792-6839 ##label BEN2  
 ##cross-references EMBL:X15423; NID:96897; PIDN:CAA33463.1; PID:96898  
 ##experimental\_source var. Bristol  
 #authors S06797  
 Benian, G.M.; Kiff, J.E.; Neckelmann, N.; Moerman, D.G.; Waterston, R.H.  
 #Journal Nature (1989) 342:45-50  
 #title Sequence of an unusually large protein implicated in regulation of myosin activity in C. elegans.  
 #cross-references MUID:9004042  
 #accession S06797  
 #status nucleic acid sequence not shown  
 ##molecule\_type DNA  
 ##residues 806-1175;1178-1998,'Y',2000-3040,'I',3042-3335,'I',3337-5693;5696-6359,'I',6361-6377;6386-6478;6541-6635;  
 6649-6742;6745-6838 ##label BEN3  
 ##cross-references EMBL:X15423  
 ##experimental\_source var. Bristol  
 #authors S57218  
 Benian, G.M.; L'Hernault, S.W.; Morris, M.E.  
 #Journal Genetics (1993) 134:1097-1104  
 #title Additional sequence complexity in the muscle gene, unc-22, and its encoded protein, twitchin, of Caenorhabditis

elegans.  
 #cross-references MUID:93387664  
 #accession SS7218  
 ##molecule\_type DNA  
 ##residues 2-99:108-194,'Q',196-206;374-468;558-753 ##label BEN4  
 ##experimental\_source var. Bristol  
 COMMENT Lack of unc-22 leads to a constant twitching of the body muscles.  
 GENETICS  
 #gene  
 #map\_position IV  
 #introns 18/3; 69/3; 143/2; 176/3; 264/2; 387/3; 413/2; 471/1; 516/3;  
 550/3; 582/3; 603/3; 628/3; 646/3; 674/3; 728/3; 767/2;  
 822/3; 897/1; 1770/1; 2115/3; 2597/2; 2651/1; 2746/1;  
 2820/3; 2948/1; 6152/3; 6691/3; 6776/1; 6808/3  
 CLASSIFICATION #superfamily twitchin; fibronectin type III repeat homology;  
 immunoglobulin homology; protein kinase homology  
 KEYWORDS ATP; autophosphorylation; duplication; muscle;  
 phosphotransferase; serine/threonine-specific protein  
 kinase

## FEATURE

806-898,899-990,  
 991-1083,1084-1175,  
 1178-1273,  
 1474-1567,  
 1770-1864,  
 2066-2158,  
 2358-2450,  
 2651-2745,  
 2948-3041,  
 3242-3336,  
 3536-3629,  
 3829-3921,  
 4124-4214,  
 4517-4611,  
 4812-4907,  
 5211-5303,  
 5304-5398,  
 5601-5693,  
 5696-5790,  
 6263-6356,  
 6386-6478,  
 6541-6635,  
 6649-6742,  
 6745-6838,  
 1274-1372,  
 1373-1473,  
 1568-1670,  
 1671-1769,  
 1865-1964,  
 1965-2065,  
 2159-2258,  
 2259-2357,  
 2451-2550,  
 2551-2650,  
 2746-2847,  
 2848-2947,  
 3042-3141,  
 3142-3241,  
 3337-3435,  
 3436-3535,  
 3630-3729,  
 3730-3828,  
 3922-4022,  
 4023-4123,  
 4215-4313,  
 4314-4415,  
 4416-4516,  
 4612-4710,  
 4711-4811,  
 4908-5009,  
 5010-5109,  
 5110-5210,  
 5399-5499,

#region motif 2\  
 5500-5600,  
 5791-5889,  
 5940-6197,  
 5948-5956,  
 5971  
 SUMMARY  
 #length 6839 #molecular-weight 753494 #checksum 1785

Query Match 82.5% Score 52; DB 2; Length 6839;  
 Best Local Similarity 66.7%; Pred. No. 2.11e-01;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 4909 KPASPOHIR 4917  
 QY 118 KPSPKHYR 126

## RESULT 2

ENTRY S26663 #type fragment  
 TITLE Microtubule-associated protein tau - human (fragment)  
 ORGANISM #formal\_name Homo sapiens #common\_name man  
 DATE 13-Jan-1995 #sequence\_revision 12-Apr-1996 #text\_change  
 08-Sep-1997  
 ACCESSIONS S26663; S41125  
 REFERENCE S26662  
 #authors Andreadis, A.; Brown, W.M.; Kosik, K.S.  
 #journal Biochemistry (1992) 31:10626-10633  
 #title Structure and novel exons of the human tau gene.  
 #cross-references MUID:93041157  
 #accession S26663  
 #status preliminary  
 ##molecule\_type DNA  
 ##residues 1-66 #label AN2  
 ##cross-references EMBL:X61371; NID:g36716; PID:g36717  
 REFERENCE S41125  
 #authors Andreadis, A.; Nissson, P.E.; Kosik, K.S.; Watkins, P.C.  
 #journal Nucleic Acids Res. (1993) 21:2217-2221  
 #title The exon trapping assay partly discriminates against  
 #cross-references MUID:93275752  
 #accession S41125  
 #status preliminary  
 ##molecule\_type DNA  
 ##residues 1-66 #label AN2

GENETICS  
 #gene GDB:MAPT; MTBT1  
 #cross-references GDB:119434; OMIM:157140  
 #map\_position 17q21-17q21  
 CLASSIFICATION #superfamily microtubule-associated protein tau; MAP2/tau  
 repeat homology  
 SUMMARY  
 #length 66 #checksum 4316

Query Match 74.6% Score 47; DB 2; Length 66;  
 Best Local Similarity 100.0%; Pred. No. 3.18e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 42 PSSPKHY 48  
 QY 119 PSSPKHY 125

## RESULT 3

ENTRY A40525 #type complete  
 TITLE proline-rich sheath protein Mf22 precursor - nematode (Brugia pahangi)  
 ORGANISM #formal\_name Brugia pahangi  
 DATE 28-Feb-1992 #sequence\_revision 10-Apr-1992 #text\_change  
 27-Jan-1995  
 ACCESSIONS A40525; S18744  
 REFERENCE A40525  
 #authors Selkirk, M.E.; Yazdanbakhsh, M.; Freedman, D.; Blaxter, M.L.;  
 Cookson, E.; Jenkins, R.E.; Williams, S.A.  
 #journal J. Biol. Chem. (1991) 266:11002-11008  
 #title A proline-rich structural protein of the surface sheath of

```

larval Brugia filarial nematode parasites.
#cross-references MUID:91250404
#accession A40525
##status preliminary
##molecule_type DNA
##residues 1-205 ##label SEL
##cross-references GB:X58063
##note the authors translated the codon GCA for residue 23 as
Arg, CIG for residue 146 as Val, and CCG for residue
180 as Pro
SUMMARY #length 205 #molecular-weight 22154 #checksum 835
Query Match 74.6%; Score 47; DB 2; Length 205;
Best Local Similarity 62.5%; Pred. No. 3.18e+00;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 175 KPTAPRHV 182
QY 118 KPSSPKHV 125
||||:|
|:::|:|

RESULT 4
ENTRY S26854 #type complete
TITLE microfilarial sheath protein, major component - nematode
ORGANISM (Brugia pahangi)
#formal_name Brugia pahangi
#sequence_revision 13-Jan-1995 #text_change
DATE 09-Sep-1997
ACCESSIONS S26854
REFERENCE S26854
#authors Selkirk, M.
#submission submitted to the EMBL Data Library, February 1991
#accession S26854
##status preliminary
##molecule_type DNA
##residues 1-205 ##label SEL
##cross-references EMBL:X58063; NID:g5951; PID:g5952
GENETICS 123/2
#introns #length 205 #molecular-weight 22199 #checksum 1003
SUMMARY
Query Match 74.6%; Score 47; DB 2; Length 205;
Best Local Similarity 62.5%; Pred. No. 3.18e+00;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 175 KPTAPRHV 182
QY 118 KPSSPKHV 125
||||:|
|:::|:|

RESULT 5
ENTRY JH0594 #type complete
TITLE vasoactive intestinal peptide receptor precursor - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change
26-Aug-1999
ACCESSIONS JH0594; S56014
REFERENCE JH0594
#authors Ishihara, T.; Shigemoto, R.; Mori, K.; Takahashi, K.; Nagata,
S.
#journal Neuron (1992) 8:811-819
#title Functional expression and tissue distribution of a novel
#cross-references MUID:92232309
#accession JH0594
##molecule_type mRNA
##residues 1-459 ##label ISH
##cross-references GB:M86835; NID:g207640; PID:g207641
##experimental_source lung
#accession S56014
#authors Pei, L.; Melmed, S.
#journal Biochem. J. (1995) 308:719-723
#title Characterization of the rat vasoactive intestinal polypeptide

```

```

receptor gene 5' region.
#cross-references MUID:97104266
#accession S56014
##status preliminary; translation not shown
##molecule_type DNA
##residues 1-26 ##label PEI
##cross-references EMBL:U10635; NID:g505752; PID:g514311
CLASSIFICATION #superfamily glucagon receptor
KEYWORDS G protein-coupled receptor; glycoprotein; intestine;
transmembrane protein
FEATURE
1-30 #domain signal sequence #status predicted #label sig\
31-459 #product vasoactive intestinal polypeptide receptor
#status predicted #label MAT\
146-168 #domain transmembrane #status predicted #label TM1\
176-195 #domain transmembrane #status predicted #label TM2\
218-241 #domain transmembrane #status predicted #label TM3\
256-277 #domain transmembrane #status predicted #label TM4\
295-318 #domain transmembrane #status predicted #label TM5\
344-363 #domain transmembrane #status predicted #label TM6\
376-395 #domain transmembrane #status predicted #label TM7\
58,69,100,292 #binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY #length 459 #molecular-weight 52057 #checksum 2598
Query Match 74.6%; Score 47; DB 2; Length 459;
Best Local Similarity 66.7%; Pred. No. 3.18e+00;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 2 RPPSPPHV 10
QY 118 KPSSPKHV 126
||||:|
|:::|:|

RESULT 6
ENTRY B70160 #type complete
TITLE ribosomal protein L22 (rplV) - Lyme disease spirochete
ORGANISM #formal_name Borrelia burgdorferi #common_name Lyme disease
spirochete
DATE 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change
13-Aug-1999
ACCESSIONS B70160
REFERENCE A70100
#authors Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.;
Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.;
Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb,
J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.;
Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.;
Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman,
J.; Otterback, T.; Watthey, L.; McDonald, L.; Artiach, P.;
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst,
K.; Roberts, K.; Hatch, B.; Smith, H.O.; Venter, J.C.
#journal Nature (1997) 390:580-586
#title Genomic sequence of a Lyme disease spirochaete,
Borrelia burgdorferi
#cross-references MUID:98065943
#accession B70160
#status preliminary; nucleic acid sequence not shown;
translation not shown
##molecule_type DNA
##residues 1-120 ##label KLE
##cross-references GB:AE001152; GB:AE000783; NID:g2688387;
PIDN:AAC66859.1; PID:g2688408; TIGR:BB0483
##experimental_source strain B31
CLASSIFICATION #superfamily Escherichia coli ribosomal protein L22
SUMMARY #length 120 #molecular-weight 13666 #checksum 3381
Query Match 73.0%; Score 46; DB 2; Length 120;
Best Local Similarity 87.5%; Pred. No. 5.36e+00;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 15 PSSPKV 22
||||:|
|:::|:|

```

```

QY 119 PSSPKHVR 126

RESULT 7
ENTRY #type fragment
TITLE microtubule-associated protein 19B - Pacific electric ray
ALTERNATE_NAMES (fragment)
ORGANISM neuron-specific protein 19B
#formal_name Torpedo californica #common_name Pacific
#electric ray
DATE 02-Feb-1990 #sequence_revision 11-Sep-1992 #text_change
13-Aug-1999
ACCESSIONS B33319
REFERENCE A33319
#authors Ngsee, J.K.; Scheller, R.H.
#journal DNA (1989) 8:555-561
#title Isolation and characterization of two homologous cDNA clones
#cross-references MUID:90091742
#accession B33319
##status Preliminary
##molecule_type mRNA
##residues 1-350 #label NGS
##cross-references GB:M30271; NID:g213236; PIDN:AAA49280.1; PID:g213237
##note the authors translated the codon CTG for residue 18 as
Val, CCC for residue 289 as Val, CAC for residue 291
as Gln, and CCC for residue 298 as Gly
CLASSIFICATION #superfamily microtubule-associated protein MAP1B
KEYWORDS microtubule binding; phosphoprotein
SUMMARY #length 350 #checksum 532

Query Match 73.0%; Score 46; DB 2; Length 350;
Best Local Similarity 66.7%; Pred. No. 5.36e+00;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 323 KPSAPKVR 331
||:|:|
QY 118 KPSSPKHVR 126

RESULT 8
ENTRY #type complete
TITLE microtubule-associated protein - fruit fly (Drosophila
melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change
20-Sep-1999
ACCESSIONS T13845
REFERENCE Z17792
#authors Saunders, R.D.; Avides, M.C.; Howard, T.; Gonzalez, C.;
Glover, D.M.
#journal J. Cell Biol. (1997) 137:881-890
#title The Drosophila gene abnormal spindle encodes a
microtubule-associated protein that associates with the
polar regions of the mitotic spindle.
#accession T13845
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-1861 #label SAU
##cross-references EMBL:U95171; NID:g1930121; PID:g1930122;
PIDN:AAB51540.1

GENETICS asp
FUNCTION is required for the normal function of the mitotic spindle
#description #length 1861 #molecular-weight 219558 #checksum 5769
SUMMARY

Query Match 73.0%; Score 46; DB 2; Length 1861;
Best Local Similarity 55.6%; Pred. No. 5.36e+00;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 471 KPATPKVR 479
||:|:|

QY 118 KPSSPKHVR 126

RESULT 9
ENTRY #type complete
TITLE hypothetical protein APE0469 - Aeropyrum pernix (strain K1)
ORGANISM #formal_name Aeropyrum pernix
DATE 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change
20-Aug-1999
ACCESSIONS D72742
REFERENCE A72450
#authors Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.;
Haikawa, Y.; Jin-no, K.; Takahashi, M.; Sekine, M.; Baba,
S.; Ankai, A.; Kosugi, H.; Hosoyama, A.; Fukui, S.; Nagai,
Y.; Nishijima, K.; Nakazawa, H.; Takamiya, M.; Masuda, S.;
Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.;
Kushida, N.; Oguchi, A.; Aoki, K.; Kubota, K.; Nakamura,
Y.; Nomura, N.; Sako, Y.; Kikuchi, H.
#journal DNA Res. (1999) 6:83-101
#title Complete genome sequence of an aerobic hyper-thermophilic
Crenarchaeon, Aeropyrum pernix K1.
#cross-references MUID:99310339
#accession D72742
##status preliminary
##molecule_type DNA
##residues 1-144 #label KAW
##cross-references DDBJ:AP000059; NID:g5103911; PIDN:BAA79432.1;
PID:dl043218; PID:g5104116
##experimental_source strain K1

GENETICS APE0469
SUMMARY #length 144 #molecular-weight 15914 #checksum 8138

Query Match 71.4%; Score 45; DB 2; Length 144;
Best Local Similarity 66.7%; Pred. No. 8.96e+00;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 17 KPASAKVR 25
||:|:|
QY 118 KPSSPKHVR 126

RESULT 10
ENTRY #type fragment
TITLE ferredoxin--NADP+ reductase (EC 1.18.1.2) - garden pea
(fragment)
ORGANISM #formal_name Pisum sativum #common_name garden pea
DATE 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change
18-Jun-1999
ACCESSIONS T06773
REFERENCE Z15802
#authors Bowsher, C.G.
#submission submitted to the EMBL Data Library, July 1996
#accession T06773
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-378 #label BOW
##cross-references EMBL:X99419; NID:e999432; PIDN:CAA67796.1;
PID:e254893

FUNCTION #description catalyzes the reversible reduction of NADP+ by reduced
ferredoxin or reduced flavodoxin
CLASSIFICATION #superfamily ferredoxin--NADP+ reductase; cytochrome-b5
reductase homology
KEYWORDS electron transfer; FAD; flavoprotein; NADP; oxidoreductase
SUMMARY #length 378 #checksum 7120

Query Match 71.4%; Score 45; DB 2; Length 378;
Best Local Similarity 66.7%; Pred. No. 8.96e+00;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 145 KPSSPKHVR 153
||:|:|

```

```
QY 118 KPSPKXVR 126

RESULT 11
ENTRY [RNA-polymerase]-subunit kinase (EC 2.7.1.141) alpha chain -
TITLE yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES CTD kinase alpha chain; CTD kinase largest chain; protein
                  YKL139w
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
ACCESSIONS S32593; S37968; S27423
REFERENCE S32593
#authors Lee, J.M.; Greenleaf, A.L.
#journal Gene Expr. (1991) 1:149-167
#title CTD kinase large subunit is encoded by CTK1, a gene required
        for normal growth of Saccharomyces cerevisiae.
#cross-references MUID:92314702
#accession S32593
#molecule_type DNA
#residues 1-528 #label LEF
#cross-references EMBL:M69024; NID:g171327; PIDN:AC41642.1;
        PID:g171328
REFERENCE S37968
#authors Ramezani Rad, M.; Xu, G.; Kirchrath, L.; Fritz, C.; Keuchel,
        H.; Hollenberg, C.P.
#submission submitted to the Protein Sequence Database, March 1994
#accession S37968
#molecule_type DNA
#residues 1-528 #label RAM
#cross-references EMBL:Z28139; NID:g486234; PIDN:CAA81980.1;
        PID:g486235; MIPS:YKL139w
GENETICS
#gene SGD:CTK1
#map_position 11
#CLASSIFICATION superfamily unassigned Ser/Thr or Tyr-specific protein
        kinases; protein kinase homology
KEYWORDS ATP; autophosphorylation; nucleus; phosphoprotein;
        phosphotransferase; serine/threonine-specific protein
        kinase
FEATURE 181-436 #domain protein kinase homology #label KIN\
186-197 #region protein kinase ATP-binding motif\
306 #active_site Asp #status predicted
SUMMARY #length 528 #molecular-weight 60501 #checksum 6953
Query Match 71.43; Score 45; DB 2; Length 528;
Best Local Similarity 44.43; P-Val. No. 8.96e+00;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 34 RPPPKRIR 42
:||||:
QY 118 KPSPKXVR 126

RESULT 12
ENTRY A49676 #type complete
TITLE nitric-oxide synthase (EC 1.14.13.39), inducible - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change
ACCESSIONS A49676; JX0345; G01947; I38933; S47566; A47475
REFERENCE A49676
#authors Charles, I.G.; Palmer, R.M.; Hickery, M.S.; Bayliss, M.T.;
        Chubb, A.P.; Hall, V.S.; Moss, D.W.; Moncada, S.
#journal Proc. Natl. Acad. Sci. U.S.A. (1993) 90:11419-11423
#title Cloning, characterization, and expression of a cDNA encoding
        an inducible nitric oxide synthase from the human
        chondrocyte.
#cross-references MUID:94068614
#accession A49676

#status preliminary
#molecule_type mRNA
#residues 1-675, 'I', 677-932, 'G', 934-965, 'A', 967-1153 #label RE2
#cross-references EMBL:U20141; NID:g687680; PIDN:AAB60386.1;
        PID:g687681
REFERENCE S47566
#authors Maier, R.; Bilbe, G.; Rediske, J.; Lotz, M.
#journal Biochim. Biophys. Acta (1994) 1208:145-150
#title Inducible nitric oxide synthase from human articular
        chondrocytes: cDNA cloning and analysis of mRNA expression.
#cross-references MUID:94368816
#accession S47566
#status preliminary; nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-22, 'G', 24-153, 'I', 155-176, 'V', 178-799, 'A', 801-912, 'P',
        914-1153 #label MAI
#cross-references EMBL:U05810; NID:g452487; PIDN:AAA56666.1;
        PID:g452488
REFERENCE A47475
#authors Geller, D.A.; Lowenstein, C.J.; Shapiro, R.A.; Nussler, A.K.;
        Di Silvio, M.; Wang, S.C.; Nakayama, D.K.; Simmons, R.L.;
        Snyder, S.H.; Billiar, T.R.
#journal Proc. Natl. Acad. Sci. U.S.A. (1993) 90:3491-3495
#title Molecular cloning and expression of inducible nitric oxide
        synthase from human hepatocytes.
#cross-references MUID:93234523
#accession A47475
#status preliminary; not compared with conceptual translation
#molecule_type mRNA
#residues 1-422, 'I', 424-804, 'D', 806-830, 'SP', 833-932, 'G', 934-965,
        'A', 967-986, 'V', 988-1153 #label GEL
#cross-references GB:L09210; NID:g292241; PIDN:AAA59171.1; PID:g292242
#experimental_source hepatocytes
#note #sequence extracted from NCBI backbone (NCBIP:129733)
GENETICS
#gene GDS:NOS2A; NOS2; INOS
#cross-references GDB:139215; OMIM:163730
```

```

#map_position 17cen-17q11.2
FUNCTION
#description catalyzes the oxidation of an L-arginine guanidino nitrogen
and of NADPH by dioxygen to produce nitric oxide,
citruilline and NADP+
CLASSIFICATION #superfamily nitric-oxide synthase; flavodoxin homology;
NADPH--ferrihemoprotein reductase homology
KEYWORDS calmodulin binding; chromoprotein; FAD; flavoprotein; FMN;
heme; iron; NADP; oxidoreductase
FEATURE
509-529 #region calmodulin binding #status predicted\
539-1127 #domain NADPH--ferrihemoprotein reductase homology
#label FEH\
541-677 #domain flavodoxin homology #label FLX\
623-654 #region FMN binding #status predicted\
765-778 #region FMN-binding phosphate binding #status predicted\
903-913 #region FAD-isoalloxazine binding #status predicted\
978-996 #region NADP-ribose binding #status predicted\
1076-1091 #region NADP-adenine binding #status predicted\
200 #binding-site heme iron (Cys) (axial ligand) #status
predicted\
SUMMARY #length 1153 #molecular-weight 131116 #checksum 4408
Query Match 71.4%; Score 45; DB 2; Length 1153;
Best Local Similarity 85.7%; Pred. No. 8.96e+00;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 80 SSPRHR 86
QY 120 SSPKVR 126
|||||
RESULT 13
ENTRY
TITLE T04703 #type complete
ALTERNATE_NAMES hypothetical protein F4B14.200 - Arabidopsis thaliana
ORGANISM hypothetical protein T19K4.60
#formal_name Arabidopsis thaliana #common_name mouse-ear
cress
DATE 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change
21-May-1999
ACCESSIONS T04703; T05485
REFERENCE
#authors Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hoheisel, J.;
Mewes, H.W.; Mayer, K.F.X.; Schueller, C.
#submission submitted to the Protein Sequence Database, October 1998
#accession T04703
##molecule_type DNA
##residues 1-296 #label BEV
##cross-references EMBL:AL031986
##experimental_source cultivar Columbia; BAC clone F4B14
REFERENCE 215418
#authors Bevan, M.; Wedler, H.; Wambutt, R.; Hoheisel, J.; Mewes,
H.W.; Mayer, K.F.X.; Schueller, C.
#submission submitted to the Protein Sequence Database, April 1998
#accession T05485
##molecule_type DNA
##residues 1-296 #label BEW
##cross-references EMBL:AL022373
##experimental_source cultivar Columbia; BAC clone T19K4
GENETICS
#map_position 4
#intron 170/3; 207/2
#note F4B14.200; T19K4.60
SUMMARY #length 296 #molecular-weight 33332 #checksum 5628
Query Match 69.8%; Score 44; DB 2; Length 296;
Best Local Similarity 66.7%; Pred. No. 1.49e-01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 270 KPMAPKPR 278
QY 118 KPSSPKHR 126
|||||

```

```

RESULT 14
ENTRY
TITLE S53305 #type fragment
ferredoxin--NADP+ reductase (EC 1.18.1.2) precursor, root -
maize (fragment)
ORGANISM #formal_name Zea mays #common_name maize
DATE 14-Jul-1995 #sequence_revision 15-May-1998 #text_change
18-Jun-1999
ACCESSIONS S53305
REFERENCE
#authors Ritchie, S.W.; Redinbaugh, M.G.; Shiraishi, N.; Vrba, J.M.;
Campbell, W.H.
#journal Plant Mol. Biol. (1994) 26:679-690
#title Identification of a maize root transcript expressed in the
primary response to nitrate: characterization of a cDNA
with homology to ferredoxin-NADP(+) oxidoreductase.
#cross-references MUID:95036048
#accession S53305
##status nucleic acid sequence not shown
##molecule_type mRNA
##residues 1-327 #label RIT
##cross-references EMBL:U10418; NID:g500750; PIDN:AAB40034.1;
PID:g500751
##experimental_source root; seedling; strain W64AXW182E
GENETICS
#genome nuclear
#complex monomer
#function probably catalyzes NADPH-dependent reduction of ferredoxin
required for assimilation of nitrite into amino acids
#pathway nitrate assimilation
#note transcription induced by nitrate
CLASSIFICATION #superfamily ferredoxin--NADP+ reductase; cytochrome-b5
electron homology
KEYWORDS electron transfer; FAD; flavoprotein; NADP; oxidoreductase;
plastid
FEATURE
1-11 #domain transit peptide (plastid) (fragment) #status
predicted #label TNP\
12-327 #product ferredoxin--NADP+ reductase #status predicted
#label MAT\
30-175 #domain FAD binding #status predicted #label FAD\
52-312 #domain cytochrome-b5 reductase homology #label CBR\
176-327 #domain NADP binding #status predicted #label NADP
SUMMARY #length 327 #checksum 2639
Query Match 69.8%; Score 44; DB 2; Length 327;
Best Local Similarity 55.6%; Pred. No. 1.49e+01;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 94 KPGAPQNR 102
QY 118 KPSSPKHR 126
|||||
RESULT 15
ENTRY
TITLE I48691 #type complete
regulatory protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
27-Feb-1997
ACCESSIONS I48691; I84715
REFERENCE
#authors Galiana, E.; Bernard, R.; Borde, I.; Rouget, P.; Evraud, C.
#journal J. Neurosci. Res. (1993) 36:133-146
#title Proliferation and differentiation properties of bipotent
glial progenitor cell lines immortalized with the
adenovirus E1A gene.
#cross-references MUID:94087786
#accession I48691
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-332 #label RES

```



US-03-376-430-2-15.rpr

Thu May 11 06:49:58 2000

```

##cross-references EMBL:X67209; NID:g452275; PID:g452276
#accession I94715
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-332 #label RE2
##cross-references GB:L03814; NID:g456106; PID:g456107
GENETICS
#gene
SUMMARY
    npdc-1
    #length 332 #molecular-weight 35804 #checksum 478
    Query Match      69.8%; Score 44; DB 2; Length 332;
    Best Local Similarity 62.5%; Pred. No. 1.49e+01;
    Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
    Db 7 PPSRHLR 14
    QY 119 PSSPKHVR 126
    Search completed: Wed May 10 13:06:50 2000
    Job time : 8 secs.

```



\*\*\*\*\*  
W P I S R L H  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 13:00:29 2000; MasPar time 89.70 Seconds  
Tabular output not generated. 3.056 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (J18-126) from US09376430A.pap (15 of 25)  
Perfect Score: 63  
Sequence: 1 KPSSPKHVR 9

Scoring table: PAM 150  
Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot38  
1:swissprot

Statistics: Mean 21.527; Variance 21.445; scale 1.004

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	pred. No.
1	47	74.6	459	1 VIPR_RAT	VASOACTIVE INTESTINAL	9.50e+01
2	46	73.0	120	1 RL22_BORBU	50S RIBOSOMAL PROTEIN	1.69e+00
3	46	73.0	350	1 ENP2_TORCA	ELECTROMOTOR NEURON-AS	1.69e+00
4	45	71.4	321	1 YAJ0_RHISN	HYPOTHETICAL 36.1 KD P	2.98e+00
5	45	71.4	377	1 FENS_PEA	FERREDOXIN--NADP REDUC	2.98e+00
6	45	71.4	528	1 CKR1_YEAST	CTD KINASE ALPHA SUBUN	2.98e+00
7	45	71.4	895	1 DYSTROGLYC	DYSTROGLYCAN PRECURSOR	2.98e+00
8	45	71.4	1153	1 NITRIC_OXIDE	NITRIC OXIDE SYNTHASE,	2.98e+00
9	44	69.8	332	1 NEURAL_PROLIF	NEURAL PROLIFERATION D	5.20e+00
10	44	69.8	590	1 OVARIAN_ABUN	OVARIAN ABUNDANT MESSA	5.20e+00
11	44	69.8	895	1 DYSTROGLYC	DYSTROGLYCAN PRECURSOR	5.20e+00
12	44	69.8	1165	1 POL_GALV	POL POLYPROTEIN [CONTA	5.20e+00
13	43	68.3	111	1 NADH-UBIQUIN	NADH-UBIQUINONE OXIDOR	8.99e+00
14	43	68.3	263	1 PYRUVATE_FOR	PYRUVATE FORMATE-LYASE	8.99e+00
15	43	68.3	263	1 A103_SCHNA	ANTIGEN 10-3 PRECURSOR	8.99e+00
16	43	68.3	378	1 FENS_ORYSA	FERREDOXIN--NADP REDUC	8.99e+00
17	43	68.3	536	1 SYEM_YEAST	GLUTAMYL-TRNA SYNTHETA	8.99e+00
18	43	68.3	610	1 CD11_ARATH	CALCIUM-DEPENDENT PROT	8.99e+00
19	43	68.3	692	1 MR11_YEAST	DOUBLE-STRAND BREAK RE	8.99e+00
20	43	68.3	1169	1 SUPPRESSOR	SUPPRESSOR OF VARIEGAT	8.99e+00
21	43	68.3	1534	1 YK16_ARATH	DNA (CYTOSINE-5)-METH	8.99e+00
22	42	66.7	127	1 HYPOTHETICAL	HYPOTHETICAL 13.9 KD P	1.54e+01
23	42	66.7	375	1 FENS_TOBAC	FERREDOXIN--NADP REDUC	1.54e+01

RESULT ID	VIPR_RAT	STANDARD	PRT	459 AA	ALIGNMENTS
AC	P30083				
DT	01-APR-1993 (Rel. 25, Created)				
DT	01-APR-1993 (Rel. 25, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR 1 PRECURSOR (VIP-R-1)				
DE	(PULINARY ADENYLYATE CYCLASE ACTIVATING POLYPEPTIDE TYPE II RECEPTOR)				
DE	(PACAP TYPE II RECEPTOR) (PACAP-R-2).				
GN	VIPRI.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
OC	Eutheria; Rodentia; Sciurpognathl; Muridae; Murinae; Rattus.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=LUNG;				
RC	MEDLINE; 92232309.				
RA	Ishihara T., Shigemoto R., Mori K., Takahashi K., Nagata S.;				
RT	"Functional expression and tissue distribution of a novel receptor for vasoactive intestinal polypeptide.";				
RL	Neuron 8:811-819(1992).				
CC	-!- FUNCTION: THIS IS A RECEPTOR FOR VIP. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE.				
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.				
CC	-!- TISSUE SPECIFICITY: IN LIVER, LUNG, INTESTINES, THYMUS AND BRAIN (MOSTLY IN THE CEREBRAL CORTEX AND HIPPOCAMPUS).				
CC	-!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL; M86835; AAA42331.1;				
DR	PIR; JH0594; JH0594.				
DR	GCRDB; GCR_0369;				
DR	PRINTS; PR00249; GPCRSECRETIN.				
DR	PRINTS; PR00491; VASOACTIVEIPR.				
DR	PRINTS; PR01154; VIP1RECEPTOR.				
DR	PROSITE; PS00649; G_PROTEIN_RECEP_F2.1; 1.				
DR	PROSITE; PS00650; G_PROTEIN_RECEP_F2.2; 1.				
DR	PFAM; PF00002; 7tm.2; 1.				
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.				





DR SGD: L0000432; CTK1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PFAM; PF00069; pkinase; 1.  
 KW transferase; Serine/threonine-protein kinase; ATP-binding;  
 KW Nuclear protein; Phosphorylation.  
 FT DOMAIN 37 44 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 183 469 PROTEIN KINASE.  
 FT NP\_BIND 189 197 ATP (BY SIMILARITY).  
 FT BINDING 212 212 ATP (BY SIMILARITY).  
 FT ACT\_SITE 306 306 BY SIMILARITY.  
 FT DOMAIN 506 528 ASN/ASP-RICH.  
 FT MOD\_RES 14 14 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 SQ SEQUENCE 528 AA; 60501 MW; 9862EB10FD476F6B CRC64;

Query Match 71.4%; Score 45; DB 1; Length 528;  
 Best Local Similarity 44.4%; Pred. No. 2.98e+00;  
 Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 34 RPPPPKRR 42  
 QY 118 KPSSPKHVR 126

RESULT 7  
 ID DAGL\_BOVIN STANDARD; PRT; 895 AA.  
 AC Q18738;  
 DT 15-FEB-2000 (Rel. 39, Created)  
 DT 15-FEB-2000 (Rel. 39, Last sequence update)  
 DT 15-FEB-2000 (Rel. 39, Last annotation update)  
 DE DYSTROGLYCAN PRECURSOR (DYSTROPHIN-ASSOCIATED GLYCOPROTEIN 1)  
 DE [CONTAINS: ALPHA-DYSTROGLYCAN (ALPHA-DG); BETA-DYSTROGLYCAN (BETA-DG)].  
 GN DAGL.

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.

[1]  
 RN Shimizu H.;  
 RA Submitted (NOV-1987) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RN CHARACTERIZATION.

RX MEDLINE; 99175209.  
 RA Saito F., Masaki T., Kamakura K., Anderson L.V.B., Fujita S.,  
 RA Fukuta-Ohi H., Sunada Y., Shimizu T., Matsumura K.;  
 RT "Characterization of the transmembrane molecular architecture of the  
 dystroglycan complex in schwann cells.";  
 RL J. Biol. Chem. 274:8240-8246(1999).

CC -!- FUNCTION: FORMS PART OF THE DYSTROPHIN-ASSOCIATED PROTEIN COMPLEX  
 (DAPC) WHICH MAY LINK THE CYTOSKELETON TO THE EXTRACELLULAR  
 MATRIX. ALPHA-DYSTROGLYCAN FUNCTIONS AS A LAMININ RECEPTOR. BINDS  
 TO SEVERAL TYPES OF ARENAVIRUSES. IS A TARGET FOR THE ENTRY OF  
 MYOBACTERIUM LEPTAE INTO PERIPHERAL NERVE SCHWANN CELLS.  
 CC -!- SUBCELLULAR LOCATION: ALPHA-DYSTROGLYCAN IS A EXTRACELLULAR  
 PROTEIN WHILE BETA-DYSTROGLYCAN IS A TYPE-1 MEMBRANE PROTEIN.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AB009079; BAA23650.1; -  
 KW Signal; Glycoprotein; Transmembrane; Cytoskeleton.  
 FT SIGNAL 1 29 POTENTIAL.  
 FT CHAIN 30 653 ALPHA-DYSTROGLYCAN.  
 FT CHAIN 654 895 BETA-DYSTROGLYCAN.  
 FT DOMAIN 654 749 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 750 775 POTENTIAL.  
 FT DOMAIN 776 895 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 182 264 BY SIMILARITY.  
 FT DISULFID 669 713 POTENTIAL.  
 FT CARBOHYD 141 141 POTENTIAL.  
 FT CARBOHYD 641 641 POTENTIAL.  
 FT CARBOHYD 649 649 POTENTIAL.  
 FT CARBOHYD 661 661 POTENTIAL.  
 SQ SEQUENCE 895 AA; 97321 MW; 400213A299630D11 CRC64;

Query Match 71.4%; Score 45; DB 1; Length 895;  
 Best Local Similarity 66.7%; Pred. No. 2.98e+00;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 303 KPSPKRR 311  
 QY 118 KPSSPKHVR 126

RESULT 8  
 ID NOS2\_HUMAN STANDARD; PRT; 1153 AA.  
 AC P35228; Q16692;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE NITRIC OXIDE SYNTHASE, INDUCIBLE (EC 1.14.13.39) (NOS, TYPE II)  
 DE (INDUCIBLE NOS) (INOS) (HEPATOXYTE NOS) (HEP-NOS).  
 GN NOS2A OR NOS2.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;  
 RX MEDLINE; 93234523.  
 RA Geller D.A., Lowenstein C.J., Shapiro R.A., Nussler A.K.,  
 RA di Silvio M., Wang S.C., Nakayama D.K., Simmons R.L., Snyder S.H.,  
 RA Billiar T.R.;  
 RT "Molecular cloning and expression of inducible nitric oxide synthase  
 from human hepatocytes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:3491-3495(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RA Sherman P.A., Laubach V.E., Reep B.R., Wood E.R.;  
 RA Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.

RX MEDLINE; 94068614.  
 RA Charles I.G., Palmer R.M.J., Hickery M.S., Bayliss M.T.,  
 RA Chubb A.P., Hall V.S., Moss D.W., Moncada S.;  
 RT "Cloning, characterization, and expression of a cDNA encoding an  
 inducible nitric oxide synthase from the human chondrocyte.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:11419-11423(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A.

RX TISSUE=ARTICULAR CHONDROCYTES;  
 RC MEDLINE; 94368816.

RA Maier R., Bilbe G., Rediske J., Lotz M.;  
 RT "Inducible nitric oxide synthase from human articular chondrocytes:  
 RT cDNA cloning and analysis of mRNA expression.";  
 RL Biochim. Biophys. Acta 1208:145-150(1994).  
 RN [5]  
 RP SEQUENCE FROM N.A.

RA Park C., Gianotti C., Park R., Krishna G.;  
 RA Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.

RC TISSUE=GLIOBLASTOMA;  
 RX MEDLINE; 95155267.

RA Hokari A., Zeniya M., Esumi H.;  
 RT "Cloning and functional expression of human inducible nitric oxide  
 RT synthase (NOS) cDNA from a glioblastoma cell line A-172.";  
 RL J. Biochem. 116:575-581(1994).

Thu May 11 06:49:59 2000

```

RN RP SEQUENCE FROM N.A.
RC TISSUE-AIRWAY EPITHELIUM;
RX MEDLINE; 95372368.
RA Guo F.H., de Raave R.H., Rice T.W., Stuehr D.J., Thunnissen F.B.J.M.,
RA Erzurum S.C.;
RT "Continuous nitric oxide synthesis by inducible nitric oxide synthase
RT in normal human airway epithelium in vivo.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:7809-7813(1995).
CC -!- FUNCTION: THIS ENZYME PRODUCES NITRIC OXIDE (NO) WHICH IS A
CC MESSENGER MOLECULE WITH DIVERSE FUNCTIONS THROUGHOUT THE BODY.
CC IN MACROPHAGES, NO MEDIATES TUMORICIDAL AND BACTERICIDAL ACTIONS.
CC -!- CATALYTIC ACTIVITY: L-ARGININE + N NADPH + M O(2) - CITRULLINE +
CC NITRIC OXIDE + N NADP(+).
CC -!- COFACTOR: BINDS ONE MOLE EACH OF FAD AND FMN.
CC -!- ENZYME REGULATION: REGULATED BY CALCIUM/CALMODULIN (IN CONSTRAST
CC WITH MOUSE NOS2).
CC -!- INDUCTION: BY TREATMENT WITH ENDOTOXINS OR CYTOKINES.
CC -!- SIMILARITY: STRONG, TO OTHER NOS ISOZYMES. ALSO TO CYTOCHROME
CC P-450 REDUCTASE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L09210; AAA59171.1; -
CC EMBL; L24553; AAA36375.1; -
CC EMBL; X73029; CAA51512.1; -
CC EMBL; U05810; AAA56666.1; -
CC EMBL; U31511; AAB49041.1; -
CC EMBL; D26525; BAA05531.1; -
CC EMBL; U20141; AAB60366.1; -
CC PIR; A47475; A47475.
CC HSSP; P00388; 1AMO.
CC MIM; 163730; -
CC PFAM; PF00175; oxidored_fad; 1.
CC PFAM; PF00667; fad_binding; 1.
CC PRINFS; P00369; ELAVODOMIN.
CC PRINFS; P00371; PFNCR.
CC Oxidoreductase; NADP; FAD; FMN; Calcium-binding; Calmodulin-binding;
CC Phosphorylation.
KW DOMAIN 509 529
KW CALMODULIN-BINDING (POTENTIAL).
FT NP_BIND 623 654
FT FAD (ADP PART) (BY SIMILARITY).
FT NP_BIND 767 788
FT FAD (FLAVIN PART) (BY SIMILARITY).
FT NP_BIND 903 913
FT NADP (RIBOSE PART) (BY SIMILARITY).
FT NP_BIND 978 996
FT NADP (ADP PART) (BY SIMILARITY).
FT NP_BIND 1076 1091
FT PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT MOD_RES 234 234
FT PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT MOD_RES 578 578
FT PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT MOD_RES 892 892
FT D -> G (IN REF. 4).
FT CONFLICT 23 23
FT F -> L (IN REF. 4).
FT CONFLICT 154 154
FT F -> V (IN REF. 4).
FT CONFLICT 177 177
FT L -> I (IN REF. 2).
FT CONFLICT 423 423
FT S -> L (IN REF. 5 AND 6).
FT CONFLICT 608 608
FT T -> I (IN REF. 7).
FT CONFLICT 676 676
FT A -> A (IN REF. 4).
FT CONFLICT 800 800
FT A -> D (IN REF. 2).
FT CONFLICT 805 805
FT EL -> SF (IN REF. 2).
FT CONFLICT 831 832
FT S -> P (IN REF. 4).
FT CONFLICT 913 913
FT R -> G (IN REF. 2 AND 7).
FT CONFLICT 933 933
FT G -> A (IN REF. 2 AND 7).
FT CONFLICT 966 966
FT A -> V (IN REF. 2).
FT CONFLICT 987 987
FT SEQUENCE 1153 AA; 131117 MW; 47671E5385CB3A52 CRC64;

```

```

Query Match 71.4%; Score 45; DB 1; Length 1153;
Best Local Similarity 85.7%; Pred. No. 2.98e+00;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 80 SSPRHVR 86
QY 120 SSPKVR 126
PRT; 332 AA.

RESULT 9
ID NPDL_MOUSE STANDARD; PRT; 332 AA.
AC Q64322;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE NEURAL PROLIFERATION DIFFERENTIATION AND CONTROL PROTEIN-1 PRECURSOR
DE (NPDC-1 PROTEIN).
DE (NPDC-1 OR NPDC-1).
GN NPDC1 OR NPDC-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE; 94087786.
RA Gallana E., Bernard R., Borde I., Rouget P., Evraud C.;
RT "Proliferation and differentiation properties of bipotent glial
RT progenitor cell lines immortalized with the adenovirus E1A gene.";
RL J. Neurosci. Res. 36:133-146(1993).
CC [2]
CC SEQUENCE FROM N.A.
CC TISSUE-BRAIN;
CC MEDLINE; 95183505.
CC Gallana E., Vernier P., Dupont E., Evraud C., Rouget P.;
CC "Identification of a neural-specific cDNA, NPDC-1, able to down-
CC regulate cell proliferation and to suppress transformation.";
CC Proc. Natl. Acad. Sci. U.S.A. 92:1560-1564(1995).
CC -!- FUNCTION: SUPPRESSES ONCOGENIC TRANSFORMATION IN NEURAL AND NON-
CC NEURAL CELLS AND DOWN-REGULATES NEURAL CELL PROLIFERATION.
CC MIGHT BE INVOLVED IN TRANSCRIPTIONAL REGULATION
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE BRAIN AND NERVOUS SYSTEM. NOT
CC DETECTED IN LIVER, HEART, SKELETAL MUSCLE, SPLEEN, PANCREAS,
CC PITUITARY AND ADRENAL GLANDS. EXPRESSION INCREASES WHEN CULTURED
CC NEURAL CELLS ARE GROWTH-ARRESTED AND BEGIN TO DIFFERENTIATE.
CC -!- SIMILARITY: TO C.ELEGANS C23H4.1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X67209; CAA47648.1; -
CC EMBL; L03814; AAA39836.1; -
CC MGD; MGI:1099802; NPDC1.1
CC Signal; Transmembrane.
FT CHAIN 35 332
FT POTENTIAL NEURAL PROLIFERATION DIFFERENTIATION AND
FT CONTROL PROTEIN-1.
FT TRANSMEM 191 211
FT DOMAIN 151 173
FT PRO/SER/THR-RICH.
FT DOMAIN 234 244
FT PRO/SER/THR-RICH.
FT SEQUENCE 332 AA; 35804 MW; 26D459B9EAD63B3 CRC64;

```

```

Query Match 69.8%; Score 44; DB 1; Length 332;
Best Local Similarity 62.5%; Pred. No. 5.20e+00;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 7 PPSPHLR 14
QY 119 PSSPKVR 126
PRT; 590 AA.

```

```

RESULT 10
ID OAM_ASCSU STANDARD; PRT; 590 AA.

```

AC Q01456;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE OVARIAN ABUNDANT MESSAGE PROTEIN (OAM PROTEIN).  
 GN OAM.  
 OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Ascaridida;  
 OC Ascaridoidea; Ascarididae; Ascaris.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC TISSUE=OVARY;  
 RX MEDLINE; 93116800.  
 RA Guindl M., Cater J., Wilson B., Charib S., Bennett K.L.;  
 RT "An extremely abundant ovarian mRNA from the parasitic nematode  
 RL Ascaris lumbricoides var. suum has multiple repeat motifs.";  
 RL Mol. Biochem. Parasitol. 56:177-180(1992)  
 CC -1- TISSUE SPECIFICITY: SOMATIC OVARIAN TISSUE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; M94771; AAA73355.1; -  
 DR PIR; A48461; A48461.  
 DR S27773; S27773.  
 KW Repeat.  
 FT DOMAIN 66 185 20 X 6 AA TANDEM REPEATS, MOTIF 1.  
 FT DOMAIN 300 347 7 X 6 AA TANDEM REPEATS, MOTIF 2.  
 FT DOMAIN 348 413 11 X 6 AA TANDEM REPEATS, MOTIF 3.  
 FT DOMAIN 419 448 5 X 6 AA TANDEM REPEATS, MOTIF 4.  
 SQ SEQUENCE 590 AA; 62963 MW; 5BDB9D0691B8BEF3 CRC64;  
 Query Match 69.8%; Score 44; DB 1; Length 590;  
 Best Local Similarity 55.6%; Pred. No. 5.20e+00;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Db 15 KSSSPRRIR 23  
 QY 118 KPSSPKHVR 126  
 RESULT 11  
 ID DAG1\_HUMAN STANDARD; PRT; 895 AA.  
 AC Q14118;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DE DYSTROGLYCAN PRECURSOR (DYSTROPHIN-ASSOCIATED GLYCOPROTEIN 1)  
 DE [CONTAINS: ALPHA-DYSTROGLYCAN (ALPHA-DG); BETA-DYSTROGLYCAN (BETA-  
 DE DG)].  
 DE DAG1.  
 GN DAG1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC TISSUE=SKETAL MUSCLE;  
 RX MEDLINE; 94093553.  
 RA Ibraghimov-Beskrovnaya O., Milatovich A., Oczelik T., Yang B.,  
 RA Koenig K., Francke U., Campbell K.P.;  
 RT "Human dystroglycan: skeletal muscle cDNA, genomic structure, origin  
 RT of tissue specific isoforms and chromosomal localization.";  
 RL Hum. Mol. Genet. 2:1651-1657(1993).  
 CC -1- FUNCTION: FORMS PART OF THE DYSTROPHIN-ASSOCIATED PROTEIN COMPLEX  
 CC (DAPC) WHICH MAY LINK THE CYTOSKELETON TO THE EXTRACELLULAR  
 CC MATRIX. ALPHA-DYSTROGLYCAN FUNCTIONS AS A LAMININ RECEPTOR. BINDS  
 CC TO SEVERAL TYPES OF ARENAVIRUSES. IS A TARGET FOR THE ENTRY OF  
 CC MYCOBACTERIUM LEPRAE INTO PERIPHERAL NERVE SCHWANN CELLS.

CC -1- SUBCELLULAR LOCATION: ALPHA-DYSTROGLYCAN IS A EXTRACELLULAR  
 CC PROTEIN WHILE BETA-DYSTROGLYCAN IS A TYPE-I MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF FETAL AND ADULT  
 CC TISSUES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; L19711; AAA81779.1; -  
 DR MIM; L28239; -  
 KW Signal; Glycoprotein; Transmembrane; Cytoskeleton.  
 FT SIGNAL 1 29 POTENTIAL.  
 FT CHAIN 30 653 ALPHA-DYSTROGLYCAN.  
 FT DOMAIN 654 895 BETA-DYSTROGLYCAN.  
 FT TRANSMEM 750 775 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 776 895 POTENTIAL.  
 FT TRANSMEM 776 895 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 182 264 POTENTIAL.  
 FT CARBOHYD 669 713 POTENTIAL.  
 FT CARBOHYD 141 141 POTENTIAL.  
 FT CARBOHYD 485 485 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).  
 FT CARBOHYD 641 641 POTENTIAL.  
 FT CARBOHYD 649 649 POTENTIAL.  
 FT CARBOHYD 661 661 POTENTIAL.  
 SQ SEQUENCE 895 AA; 97580 MW; 38D6D7431DFA82DF CRC64;  
 Query Match 69.8%; Score 44; DB 1; Length 895;  
 Best Local Similarity 66.7%; Pred. No. 5.20e+00;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 303 KPPLPKRVR 311  
 QY 118 KPSSPKHVR 126  
 RESULT 12  
 ID POL\_GALV STANDARD; PRT; 1165 AA.  
 AC P21414;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE POL POLYPROTEIN [CONTAINS: PROTEASE (EC 3.4.23.-); REVERSE  
 DE TRANSCRIPTASE (EC 2.7.7.49); ENDONUCLEASE].  
 GN POL.  
 OS Gibbon ape leukemia virus.  
 OC Viruses; Retroviridae; Retroviridae; Mammalian type C retroviruses.  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX MEDLINE; 90051069.  
 RA Delassus S., Sonigo P., Wain-Hobson S.;  
 RT "Genetic organization of gibbon ape leukemia virus.";  
 RL Virology 173:205-213(1989).  
 CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A2; ALSO KNOWN AS THE  
 CC RETROPEPSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; M26927; AAA46810.1; -  
 DR PIR; B32595; GNJLGL.  
 DR HSP; P03355; LMML.  
 DR PROSITE; PS00141; ASP\_PROTEASE; 1.



Thu May 11 06:49:59 2000

```

DR PFAM; PF00075; inaseH; 1.
DR PFAM; PF00077; ivp; 1.
DR PFAM; PF00078; ivt; 1.
DR PFAM; PF00552; integrase; 1.
DR PFAM; PF00665; ive; 1.
DR PFAM; PF00665; rna-directed DNA polymerase;
KW Hydrolase; Transferase; RNA-directed DNA polymerase;
KW Aspartyl protease; Endonuclease; Polyprotein.
FT CHAIN 1 103
FT ACT_SITE 27 27 BY SIMILARITY.
FT ACT_SITE 27 27 BY SIMILARITY.
SQ SEQUENCE 1165 AA; 129886 MW; 8B7AFD54812B7E1A CRC64;

Query Match 69.8%; Score 44; DB 1; Length 1165;
Best Local Similarity 50.0%; Pred. NO. 5.20e+00;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 411 PTPRQVR 418
|:|:|:|
QY 119 PSSPKHVR 126

RESULT 13
ID NUAM_CAICR STANDARD; PRT; 111 AA.
AC Q34076;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 4 (EC 1.6.5.3) (FRAGMENT).
GN ND4.
OS Calman crocodilus (Spectacled calman) (Calman sclerops).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
OC Crocodyliidae; Alligatorinae; Calman.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96073446.
RA Kunazawa Y., Nishida M.;
RT "Variations in mitochondrial trna gene organization of reptiles as
RT phylogenetic markers."
RL Mol. Biol. Evol. 12:759-772(1995).
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; D38190; BAA07383.1;
DR PFAM; PF00361; oxidoreductase; 1.
KW OXIDOREDUCTASE; NADH; UBIQUINONE; Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 111 AA; 12070 MW; D211FBBAA07A42D83 CRC64;

Query Match 68.3%; Score 43; DB 1; Length 111;
Best Local Similarity 66.7%; Pred. NO. 8.99e+00;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 78 KPSPSQTR 86
|:|:|:|
QY 118 KPSPKHKVR 126

RESULT 14
ID PFLA_STRMU STANDARD; PRT; 263 AA.
AC O68575; 1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PYRUVATE FORMATE-LIASE ACTIVATING ENZYME (EC 1.97.1.4) (PFL-ACTIVATING
DE ENZYME).
GN ACT OR PFLC.

```

```

OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U11.
RA Boyd D.A., Hamilton I.R., Cvitkovitch D.G., Bleiweis A.S.;
RT "Defects in D-alanyl-lipoteichoic acid synthesis in Streptococcus
RT mutants leads to acid sensitivity."
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5.
RA Yamamoto Y., Sato Y., Takahashi-Abbe S., Yamada T., Kizaki H.;
RT "Cloning and characterization of the act gene encoding pyruvate
RT formate-lyase-activating enzyme from Streptococcus mutans."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATION OF PYRUVATE FORMATE-LYASE UNDER ANAEROBIC
CC CONDITIONS BY GENERATION OF AN ORGANIC FREE RADICAL. USING
CC S-ADENOSYLMETHIONINE AND REDUCED FLAVODOXIN AS COSUBSTRATES TO
CC PRODUCE 5'-DEOXY-ADENOSINE.
CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DIHYDROFLAVODOXIN +
CC [PYRUVATE FORMATE-LYASE] - GLYCINE -> 5'-DEOXYADENOSINE + METHIONINE
CC + FLAVODOXIN + [PYRUVATE FORMATE-LYASE]-GLYCINE RADICAL.
CC -!- COFACTOR: IRON-DEPENDENT.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ORGANIC RADICAL ACTIVATING ENZYMES
CC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AF051356; AAC05773.1;
DR EMBL; AB018417; BAA34998.1;
DR PROSITE; PS01087; RADICAL_ACTIVATING; 1.
KW OXIDOREDUCTASE; IRON.
FT METAL 37 37 IRON (POTENTIAL).
FT METAL 41 41 IRON (POTENTIAL).
FT METAL 44 44 IRON (POTENTIAL).
SQ SEQUENCE 263 AA; 30150 MW; 573405FEE248EFC CRC64;

Query Match 68.3%; Score 43; DB 1; Length 263;
Best Local Similarity 55.6%; Pred. NO. 8.99e+00;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 235 KPPTDPRVR 243
|:|:|:|
QY 118 KPSPKHKVR 126

RESULT 15
ID A103_SCHMA STANDARD; PRT; 263 AA.
AC P13492;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE ANTIGEN 10-3 PRECURSOR.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PUERTO RICAN;
RX MEDLINE; 89096846.
RA Davis R.E., Davis A.H., Carroll S.M., Rajkovic A., Rottman F.M.;
RT "Randomly repeated exons encode 81-base repeats in multiple,
RT developmentally regulated Schistosoma mansoni transcripts."
RL Mol. Cell. Biol. 8:4745-4755(1988).

```

```

CC  -!- ALTERNATIVE PRODUCTS: DIFFERENT-SIZED TRANSCRIPTS ARE EXPRESSED
CC  IN THE ADULT AND CERCARIAL STAGES. THESE TRANSCRIPTS APPEARED
CC  TO BE DERIVED IN PART BY DEVELOPMENTALLY CONTROLLED ALTERNATIVE
CC  SPLICING OF SMALL EXONS AND THE USE OF ALTERNATIVE TRANSCRIPTION
CC  INITIATION SITES. THESE TRANSCRIPTS ARE HIGHLY SIMILAR AND CONTAIN
CC  VARIABLE NUMBERS OF IDENTICAL DIRECT TANDEM REPEATS OF 81 BASES.
CC  -!- DISEASE: THIS ANTIGEN IS RECOGNIZED BY SERA FROM INFECTED
CC  HUMAN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; M22346; AAA29855.1; -.
CC  PIR; A31561; A31561.
CC  FT SIGNAL 1; Signal; Alternative splicing.
CC  KW Antigen; Repeat; Signal; Alternative splicing.
CC  FT CHAIN ? 263 ANTIGEN 10-3.
CC  FT CARBOHYD 23 23 POTENTIAL.
CC  FT VARIANT 61 70 MISSING (IN SOME FORMS).
CC  FT DOMAIN 81 189 5 X 27 AA TANDEM REPEATS.
CC  FT REPEAT 81 107 1.
CC  FT REPEAT 108 134 2.
CC  FT REPEAT 135 161 3.
CC  FT REPEAT 162 188 4.
CC  FT REPEAT 189 206 5.
CC  SQ SEQUENCE 263 AA; 29640 MW; A33A70AA1E23E74E CRC64;
CC
CC  Query Match 68.3%; Score 43; DB 1; Length 263;
CC  Best Local Similarity 50.0%; Pred. NO. 8.99e+00;
CC  Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
CC
CC  Db 84 KPTTPKQI 91
CC  QY 118 KPSSPKHV 125
CC
CC  Search completed: Wed May 10 13:02:08 2000
CC  Job time : 99 secs.

```

\*\*\*\*\*  
 M P S R C H  
 \*\*\*\*\*  
 (TW)

Release 3.1A John F. Collins, BioComputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 13:02:27 2000; MasPar time 225.17 Seconds  
 Tabular output not generated. 2.771 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (118-126) from US09376430A.pep (15 of 25)  
 Perfect Score: 63  
 Sequence: 1 KPSSPKHVR 9

Scoring table: PAM 150  
 Gap 11

Searched: 225878 seqs, 69334122 residues  
 Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: sptrembl12  
 1:sp-archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
 5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
 9:sp\_phase 10:sp\_plant 11:sp\_rodent 12:sp\_unclassified  
 13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 20.701; Variance 22.161; scale 0.934

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	52	82.5	6048	5	Q23020 TWITCHIN.	1.40e-01
2	52	82.5	6831	5	UNC-22 PROTEIN.	1.40e-01
3	52	82.5	7160	5	ZK617.1B PROTEIN.	1.40e-01
4	51	81.0	222	3	HYPOTHETICAL 25.0 KD P	2.48e-01
5	47	74.6	66	4	ALTERNATIVELY SPLICED MAJOR PROTEIN COMPONENT	2.48e+00
6	47	74.6	205	5	Q00032 TMTPA.	2.48e+00
7	47	74.6	221	2	Q9XCF4 RESPONSE REGULATOR PHO	4.32e+00
8	47	74.6	234	2	Q56180 32K PRECURSOR.	4.32e+00
9	46	73.0	286	14	P89028 HYPOTHETICAL 37.1 KD P	4.32e+00
10	46	73.0	340	10	Q23131 INOSITINE-GUANOSINE KINA	4.32e+00
11	46	73.0	434	2	Q56875 CYCLIN ANTA-6A	4.32e+00
12	46	73.0	531	11	Q9WV44 REGULATOR OF G-PROTEIN	4.32e+00
13	46	73.0	1175	5	Q9YX88 MICROTUBULE ASSOCIATED	4.32e+00
14	46	73.0	1861	5	Q01401 144AA LONG HYPOTHETICA	7.45e+00
15	45	71.4	144	1	Q9YBW4 INDUCIBLE NITRIC OXIDE	7.45e+00
16	45	71.4	162	6	Q46660 [NIFE] HYDROGENASE-LIKE	7.45e+00
17	45	71.4	344	2	P94154 INDUCIBLE NITRIC OXIDE	7.45e+00
18	45	71.4	1114	4	Q94994 INDUCIBLE NITRIC OXIDE	7.45e+00
19	45	71.4	1153	4	O60757 RAS-GRF2 (FRAGMENT).	7.45e+00
20	45	71.4	1237	4	O14827	7.45e+00

21	44	69.8	296	10	O65630 HYPOTHETICAL 33.3 KD P	1.28e+01
22	44	69.8	327	10	Q41736 FERREDOXIN-NADP REDUCT	1.28e+01
23	44	69.8	533	11	Q60787 76 KD TYROSINE PHOSPHO	1.28e+01
24	44	69.8	576	10	O80673 CPDK-RELATED PROTEIN.	1.28e+01
25	44	69.8	582	10	O82649 MAP3K ALPHA PROTEIN K1	1.28e+01
26	44	69.8	601	10	O04290 CDPK-RELATED PROTEIN K	1.28e+01
27	44	69.8	608	10	Q9ZRF7 MEK KINASE.	1.28e+01
28	44	69.8	653	10	Q96304 SCARECROW.	1.28e+01
29	44	69.8	791	14	Q98252 MC085L.	1.28e+01
30	44	69.8	1127	14	Q98522 POL POLYPROTEIN.	1.28e+01
31	44	69.8	1203	14	O89815 POL POLYPROTEIN (FRAGM	1.28e+01
32	43	68.3	357	4	Q75909 CYCLIN K.	2.17e+01
33	43	68.3	444	5	Q21005 F5866.4 PROTEIN.	2.17e+01
34	43	68.3	619	3	Q98234 CHROMOSOME XV READING	2.17e+01
35	43	68.3	691	3	Q19243 SIMILAR TO SER/THR PRO	2.17e+01
36	43	68.3	692	3	Q07173 MRE11P.	3.64e+01
37	42	66.7	288	1	O86132 288AA LONG HYPOTHETICA	3.64e+01
38	42	66.7	442	2	O84713 TRIGGER FACTOR-PEPTIDY	3.64e+01
39	42	66.7	474	10	O04036 HYPOTHETICAL 52.2 KD P	3.64e+01
40	42	66.7	494	1	O28800 SIGNAL-TRANSDUCING HIS	3.64e+01
41	42	66.7	1055	2	O53348 HYPOTHETICAL 110.7 KD	3.64e+01
42	42	66.7	1083	11	Q63744 RHOGAP PROTEIN.	3.64e+01
43	42	66.7	1091	4	O43199 DELETED IN LIVER CANCER	3.64e+01
44	42	66.7	1175	5	Q20642 F52B5.1 PROTEIN.	3.64e+01
45	42	66.7	1692	4	Q12914 G2 PROTEIN (FRAGMENT).	3.64e+01

## ALIGNMENTS

RESULT 1  
 ID Q23020 PRELIMINARY; PRT: 6048 AA.  
 AC Q23020: Q27232;  
 DF 01-NOV-1996 (fremblrel. 01, Created)  
 DT 01-NOV-1996 (fremblrel. 01, Last sequence update)  
 DE TWITCHIN.  
 GN UNC-22 OR ZK617.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 90044042.  
 RA BENJAN G.M., KIFF J.E., NECKELMANN N., MOERMAN D.G., WATERSON R.H.;  
 RT "Sequence of an unusually large protein implicated in regulation of  
 myosin activity in C. elegans.";  
 RL Nature 342:45-50(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 93387664.  
 RA BENJAN G.M., L'HERNAULT S.W., MORRIS M.E.;  
 RT "Additional sequence complexity in the muscle gene, unc-22, and its  
 encoded protein, twitchin, of Caenorhabditis elegans.";  
 RL Genetics 134:1097-1104(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA HARRIS B.;  
 RT Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; X15423; CAA33463.1;  
 DR EMBL; X15423; CAA33463.1; ALT\_INIT.  
 DR EMBL; X73899; CAA98081.1;  
 DR EMBL; X73899; CAA98081.1; JOINED.  
 DR HSSP; O63450; 1A06.  
 DR PFAM; PF00041; fn3; 31.  
 DR PFAM; PF00047; ig; 13.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PRINTS; PR00014; FNTYPEII.  
 KW Myosin; Kinase.  
 SQ SEQUENCE 6048 AA: 668449 MW: 1977C602 CRC32;  
 Query Match 82.5%; Score 52; DB 5; Length 6048;  
 Best Local Similarity 66.7%; Pred. No. 1.40e-01;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 4118 KPASPQHIR 4126  
QY 118 KPSSPKHVR 126

## RESULT 2

ID Q23550 PRELIMINARY; PRT; 6831 AA.  
AC Q23550;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
DE UNC-22 PROTEIN.  
GN  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.

RA WHITE S.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z73897; CAA98064.1; -  
DR EMBL; Z73899; CAA98064.1; JOINED.  
DR HSSP; P02751; 1FNA.  
DR PFAM; PF00041; fn3; 31.  
DR PFAM; PF00047; ig; 17.  
DR PFAM; PF00069; pkinase; 1.  
DR PRINTS; PR00014; FNTYPEII.  
DR SEQUENCE 6831 AA; 752579 MW; 0A66C338 CRC32;  
SQ

Query Match 82.5%; Score 52; DB 5; Length 6831;  
Best Local Similarity 66.7%; Pred. No. 1.40e-01;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 4901 KPASPQHIR 4909  
QY 118 KPSSPKHVR 126

## RESULT 3

ID Q23551 PRELIMINARY; PRT; 7160 AA.  
AC Q23551;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
DE ZK617.1B PROTEIN.  
GN ZK617.1B.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA HARRIS B.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RX MEDLINE; 94150718.  
RA WILSON R.; AINSWORTH R.; ANDERSON K.; BAYNES C.; BERKS M.;  
RA BONFIELD J.; BORTON J.; CONNELL M.; COPSEY T.; COOPER J.; COULSON A.;  
RA CRAXTON M.; DEAR S.; DU Z.; DURBIN R.; FAVELLO A.; FULTON L.;  
RA GARDNER A.; GREEN P.; HAWKINS T.; HILLIER L.; JTER M.; JOHNSTON L.;  
RA JONES M.; KERSHAW J.; KIRSTEN J.; LAISTER N.; LATREILLE P.;  
RA LIGHTNING J.; LLOYD C.; MCMURRAY A.; MORTIMORE B.; O'CALLAGHAN M.;  
RA PARSONS J.; PERCY C.; RIFKEN L.; ROOPRA A.; SAUNDERS D.; SHOWKNEEN R.;  
RA SMALDON N.; SMITH A.; SONNHAMMER E.; STADEN R.; SULSTON J.;  
RA THIRRY-MIEG J.; THOMAS K.; VAUDIN M.; VAUGHAN K.; WATERSTON R.;  
RA WATSON A.; WEINSTOCK L.; WILKINSON-SPROAT J.; WOHLMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans."  
RL Nature 368:32-38(1994).  
RN [3]  
RP SEQUENCE FROM N.A.

Query Match 82.5%; Score 52; DB 5; Length 6831;  
Best Local Similarity 66.7%; Pred. No. 1.40e-01;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 4901 KPASPQHIR 4909  
QY 118 KPSSPKHVR 126

## RESULT 3

ID Q23551 PRELIMINARY; PRT; 7160 AA.  
AC Q23551;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
DE ZK617.1B PROTEIN.  
GN ZK617.1B.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA HARRIS B.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RX MEDLINE; 94150718.  
RA WILSON R.; AINSWORTH R.; ANDERSON K.; BAYNES C.; BERKS M.;  
RA BONFIELD J.; BORTON J.; CONNELL M.; COPSEY T.; COOPER J.; COULSON A.;  
RA CRAXTON M.; DEAR S.; DU Z.; DURBIN R.; FAVELLO A.; FULTON L.;  
RA GARDNER A.; GREEN P.; HAWKINS T.; HILLIER L.; JTER M.; JOHNSTON L.;  
RA JONES M.; KERSHAW J.; KIRSTEN J.; LAISTER N.; LATREILLE P.;  
RA LIGHTNING J.; LLOYD C.; MCMURRAY A.; MORTIMORE B.; O'CALLAGHAN M.;  
RA PARSONS J.; PERCY C.; RIFKEN L.; ROOPRA A.; SAUNDERS D.; SHOWKNEEN R.;  
RA SMALDON N.; SMITH A.; SONNHAMMER E.; STADEN R.; SULSTON J.;  
RA THIRRY-MIEG J.; THOMAS K.; VAUDIN M.; VAUGHAN K.; WATERSTON R.;  
RA WATSON A.; WEINSTOCK L.; WILKINSON-SPROAT J.; WOHLMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans."  
RL Nature 368:32-38(1994).  
RN [3]  
RP SEQUENCE FROM N.A.

Query Match 81.0%; Score 51; DB 3; Length 222;  
Best Local Similarity 55.6%; Pred. No. 2.52e-01;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 152 RPTSPKQIR 160  
QY 118 KPSSPKHVR 126

## RESULT 5

ID Q15550 PRELIMINARY; PRT; 66 AA.  
AC Q15550;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1998 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 08, Last annotation update)  
DE ALTERNATIVELY SPLICED TAU (FRAGMENT).  
GN TAU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; BROWN W.M.; KOSIK K.S.;  
RA MEDLINE; 93041757.  
RT "Structure and novel exons of the human tau gene."  
RL Biochemistry 31:10626-10633(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; NISSON P.E.; KOSIK K.S.; WATKINS P.C.;

Query Match 81.0%; Score 51; DB 3; Length 222;  
Best Local Similarity 55.6%; Pred. No. 2.52e-01;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 152 RPTSPKQIR 160  
QY 118 KPSSPKHVR 126

## RESULT 5

ID Q15550 PRELIMINARY; PRT; 66 AA.  
AC Q15550;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1998 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 08, Last annotation update)  
DE ALTERNATIVELY SPLICED TAU (FRAGMENT).  
GN TAU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; BROWN W.M.; KOSIK K.S.;  
RA MEDLINE; 93041757.  
RT "Structure and novel exons of the human tau gene."  
RL Biochemistry 31:10626-10633(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; NISSON P.E.; KOSIK K.S.; WATKINS P.C.;

Query Match 81.0%; Score 51; DB 3; Length 222;  
Best Local Similarity 55.6%; Pred. No. 2.52e-01;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 152 RPTSPKQIR 160  
QY 118 KPSSPKHVR 126

## RESULT 5

ID Q15550 PRELIMINARY; PRT; 66 AA.  
AC Q15550;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1998 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 08, Last annotation update)  
DE ALTERNATIVELY SPLICED TAU (FRAGMENT).  
GN TAU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; BROWN W.M.; KOSIK K.S.;  
RA MEDLINE; 93041757.  
RT "Structure and novel exons of the human tau gene."  
RL Biochemistry 31:10626-10633(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; NISSON P.E.; KOSIK K.S.; WATKINS P.C.;

RA WHITE S.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z73899; CAA98082.1; -  
DR EMBL; Z73897; CAA98082.1; JOINED.  
DR EMBL; Z73899; CAA98085.1; -  
DR EMBL; Z73899; CAA98085.1; JOINED.  
DR HSSP; P02751; 1FNA.  
DR PFAM; PF00041; fn3; 31.  
DR PFAM; PF00047; ig; 17.  
DR PFAM; PF00069; pkinase; 1.  
DR PRINTS; PR00014; FNTYPEIII.  
DR SEQUENCE 7160 AA; 789211 MW; EDD567FE CRC32;  
SQ

Query Match 82.5%; Score 52; DB 5; Length 7160;  
Best Local Similarity 66.7%; Pred. No. 1.40e-01;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 5230 KPASPQHIR 5238  
QY 118 KPSSPKHVR 126

## RESULT 4

ID O94706 PRELIMINARY; PRT; 222 AA.  
AC O94706;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE HYPOTHETICAL 25.0 KD PROTEIN.  
GN SPC1259.06.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Archiascomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA WOOD V.; RAJANDREAM M.A.; BARRELL B.G.; BOTHE G.; POHL T.;  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL034564; CAA22544.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 222 AA; 25003 MW; E66F8C7 CRC32;  
QY

Query Match 81.0%; Score 51; DB 3; Length 222;  
Best Local Similarity 55.6%; Pred. No. 2.52e-01;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 152 RPTSPKQIR 160  
QY 118 KPSSPKHVR 126

## RESULT 5

ID Q15550 PRELIMINARY; PRT; 66 AA.  
AC Q15550;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1998 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 08, Last annotation update)  
DE ALTERNATIVELY SPLICED TAU (FRAGMENT).  
GN TAU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; BROWN W.M.; KOSIK K.S.;  
RA MEDLINE; 93041757.  
RT "Structure and novel exons of the human tau gene."  
RL Biochemistry 31:10626-10633(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; NISSON P.E.; KOSIK K.S.; WATKINS P.C.;

Query Match 81.0%; Score 51; DB 3; Length 222;  
Best Local Similarity 55.6%; Pred. No. 2.52e-01;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 152 RPTSPKQIR 160  
QY 118 KPSSPKHVR 126

## RESULT 5

ID Q15550 PRELIMINARY; PRT; 66 AA.  
AC Q15550;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1998 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 08, Last annotation update)  
DE ALTERNATIVELY SPLICED TAU (FRAGMENT).  
GN TAU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; BROWN W.M.; KOSIK K.S.;  
RA MEDLINE; 93041757.  
RT "Structure and novel exons of the human tau gene."  
RL Biochemistry 31:10626-10633(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; NISSON P.E.; KOSIK K.S.; WATKINS P.C.;

Query Match 81.0%; Score 51; DB 3; Length 222;  
Best Local Similarity 55.6%; Pred. No. 2.52e-01;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 152 RPTSPKQIR 160  
QY 118 KPSSPKHVR 126

## RESULT 5

ID Q15550 PRELIMINARY; PRT; 66 AA.  
AC Q15550;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1998 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 08, Last annotation update)  
DE ALTERNATIVELY SPLICED TAU (FRAGMENT).  
GN TAU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; BROWN W.M.; KOSIK K.S.;  
RA MEDLINE; 93041757.  
RT "Structure and novel exons of the human tau gene."  
RL Biochemistry 31:10626-10633(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; NISSON P.E.; KOSIK K.S.; WATKINS P.C.;

Query Match 81.0%; Score 51; DB 3; Length 222;  
Best Local Similarity 55.6%; Pred. No. 2.52e-01;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 152 RPTSPKQIR 160  
QY 118 KPSSPKHVR 126

## RESULT 5

ID Q15550 PRELIMINARY; PRT; 66 AA.  
AC Q15550;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1998 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 08, Last annotation update)  
DE ALTERNATIVELY SPLICED TAU (FRAGMENT).  
GN TAU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; BROWN W.M.; KOSIK K.S.;  
RA MEDLINE; 93041757.  
RT "Structure and novel exons of the human tau gene."  
RL Biochemistry 31:10626-10633(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; NISSON P.E.; KOSIK K.S.; WATKINS P.C.;

Query Match 81.0%; Score 51; DB 3; Length 222;  
Best Local Similarity 55.6%; Pred. No. 2.52e-01;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 152 RPTSPKQIR 160  
QY 118 KPSSPKHVR 126

## RESULT 5

ID Q15550 PRELIMINARY; PRT; 66 AA.  
AC Q15550;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1998 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 08, Last annotation update)  
DE ALTERNATIVELY SPLICED TAU (FRAGMENT).  
GN TAU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; BROWN W.M.; KOSIK K.S.;  
RA MEDLINE; 93041757.  
RT "Structure and novel exons of the human tau gene."  
RL Biochemistry 31:10626-10633(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA ANDREADIS A.; NISSON P.E.; KOSIK K.S.; WATKINS P.C.;

Query Match 81.0%; Score 51; DB 3; Length 222;  
Best Local Similarity 55.6%; Pred. No. 2.52e-01;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 152 RPTSPKQIR 160  
QY 118 KPSSPKHVR 126

## RESULT 5

```

RT "The exon trapping assay partly discriminates against alternatively
RL spliced exons."
RL Nucleic Acids Res. 21:2217-2221(1993).
DR EMBL; AF047859; -; NOT_ANNOTATED_CDS.
KW Alternative splicing.
FT NON_TER
SQ SEQUENCE 66 AA; 6816 MW; 56DB9D2E CRC32;

Query Match 74.6%; Score 47; DB 4; Length 66;
Best Local Similarity 100.0%; Pred. No. 2.48e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 42 PSSPKHV 48
QY 119 PSSPKHV 125

RESULT 6
ID Q00032 PRELIMINARY; PRT; 205 AA.
AC Q00032;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE MAJOR PROTEIN COMPONENT OF THE MICROFILARIAL (L1) SHEATH.
GN F22.
OS Brugia pahangi.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Spirurida;
OC Filarioidea; Onchocercidae; Brugia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91250404.
RA SELKIRK M., YAZDANBAKHSH M., FREEDMAN D., BAXTER M., COOKSON E.,
RA JENKINS R.E., WILLIAMS S.A.;
RT "A proline-rich structural protein of the surface sheath of larval
RT Brugia filarial nematode parasites."
RL J. Biol. Chem. 266:11002-11008(1991).
DR EMBL; X58063; CAA41094.1; -;
KW Structural protein.
SQ SEQUENCE 205 AA; 22199 MW; 112FC96E CRC32;

Query Match 74.6%; Score 47; DB 5; Length 205;
Best Local Similarity 62.5%; Pred. No. 2.48e+00;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 175 KPTAPRHV 182
QY 118 KPSSPKHV 125

RESULT 7
ID Q9XCF4 PRELIMINARY; PRT; 221 AA.
AC Q9XCF4;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE TMTPA.
GN TMTPA.
OS Mycobacterium avium.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=2151;
RA ECKSTEIN T.M., LAMBERT M.L., BRENNAN P.J., BELISLE J.T., INAMINE T.M.;
RT "Identification of a gene cluster involved in glycopeptidolipid
RT biosynthesis and of a gene cluster encoding daunorubicin resistance in
RT two strains of Mycobacterium avium serovar 2."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF143772; AAD44232.1; -;
SQ SEQUENCE 221 AA; 24277 MW; D6C0013E CRC32;

Query Match 74.6%; Score 47; DB 2; Length 221;
Best Local Similarity 66.7%; Pred. No. 2.48e+00;

```

```

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 127 KPENPKHV 135
QY 118 KPSSPKHV 126

RESULT 8
ID Q56180 PRELIMINARY; PRT; 234 AA.
AC Q56180;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE RESPONSE REGULATOR PHOB.
GN PHOB.
OS Synechococcus sp. (strain WH7803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=WH7803;
RA WATSON G.M.F., SCANLAN D.J., MANN N.H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U38912; AAB38749.1; -;
DR HSSP; P41789; INTR.
DR PFAM; PF00072; response_reg; 1.
DR PFAM; PF00486; trans_reg_C; 1.
SQ SEQUENCE 234 AA; 26325 MW; F2F7530B CRC32;

Query Match 74.6%; Score 47; DB 2; Length 234;
Best Local Similarity 75.0%; Pred. No. 2.48e+00;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 217 PSSPQHIR 224
QY 119 PSSPKHV 126

RESULT 9
ID P89028 PRELIMINARY; PRT; 286 AA.
AC P89028;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE 32K PRECURSOR.
GN P32.
OS Ovine adenovirus.
OC Viruses; dsDNA-viruses, no RNA stage; Adenoviridae; Adenovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=OAV287;
RA VITALI S., BROOKES D.E., STRIKE P., KHATRI A., BOYLE D.B., BOTH G.W.;
RT "Unique genome arrangement of an ovine adenovirus: identification of
RT new proteins and proteinase cleavage sites."
RL Virology 220:186-199(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=OAV287;
RA XU Z.Z., HATT A., BOYLE D.B., BOTH G.W.;
RT "Construction of ovine adenovirus recombinants by gene insertion or
RT deletion of related terminal region sequences."
RL Virology 230:62-71(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=OAV287;
RA KHATRI A., BOTH G.W.;
RT "Identification of transcripts and promoter regions of ovine
RT adenovirus OAV287."
RL Virology 245:128-141(1998).
RN [4]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-QAV287;
RA BOTH G.W.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40839; AAB19237.2; -. 32K PROTEIN.
FT CHAIN 13 286
SQ SEQUENCE 286 AA; 32163 MW; F454AA96 CRC32;

Query Match
Best Local Similarity 73.0%; Score 46; DB 14; Length 286;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 144 SSPKHIR 150
QY 120 SSPKHVR 126

RESULT 10
ID O23131 PRELIMINARY; PRT; 340 AA.
AC O23131;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE HYPOTHETICAL 37.1 KD PROTEIN.
GN F19G10.13.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA FEDERSPIEL N.A., PALM C.J., CONWAY A.B., KURTZ D.B., CONWAY A.R.,
RA AU M., ARAUJO R., BUEHLER E., DEWAR K., FENG J., KIM C., LI Y.,
RA OJI O., OSBORNE B.I., SHINN P., SUN H., TORIUMI M., VYOTSKAIA V.,
RA YU G., ECKER J., THEOLOGIS A., DAVIS R.W.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF000657; AAB72160.1; -.
DR HSP; P09651; LUPI.
DR PFAM; PF00076; rim; 1.
KW Hypothetical protein.
FT DOMAIN 126 129 POLY-GLY.
FT DOMAIN 147 154 POLY-PRO.
FT DOMAIN 204 207 POLY-GLY.
FT DOMAIN 210 215 POLY-ALA.
FT DOMAIN 260 263 POLY-LEU.
SQ SEQUENCE 340 AA; 37084 MW; 057410FA CRC32;

Query Match
Best Local Similarity 73.0%; Score 46; DB 10; Length 340;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 111 KPSTPNHV 118
QY 118 KPSSPKHV 125

RESULT 11
ID Q56875 PRELIMINARY; PRT; 434 AA.
AC Q56875; Q56858;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE INOSINE-GUANOSINE KINASE (EC 2.7.1.73).
GN GSK.
OS Versinia enterocolitica.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Versinia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-8081C / SEROTYPE O:8;
RX MEDLINE; 97086507.
RA ZHANG L., TOIVANEN P., SKURNIK M.;

```

```

RT "The gene cluster directing O-antigen biosynthesis in Versinia
RT enterocolitica serotype O:8: identification of the genes for mannose
RT and galactose biosynthesis and the gene for the O-antigen
RT polymerase.";
RL Microbiology 142:277-288(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-8081C / SEROTYPE O:8;
RA PIERSON D.E., CARLSON S.;
RT Identification of the gale gene and a gale homolog and
RT characterization of their roles in the biosynthesis of
RT lipopolysaccharide in a serotype O:8 strain of Versinia
RT enterocolitica.";
RL J. Bacteriol. 178:5916-5924(1996).
CC -I- CATALYTIC ACTIVITY: ATP + INOSINE = ADP + IMP.
CC -I- SIMILARITY: BELONGS TO A FAMILY OF CARBOHYDRATE KINASES THAT
CC -I- GROUPS TOGETHER PFKB, FRUK, GSK, LACC, RBSK, AND SCRK.
DR EMBL; U46859; AAC60779.1; -.
DR EMBL; U43708; AAC44472.1; -.
DR PROSITE; PS00583; PFKB_KINASES.1; FALSE_NEG.
DR PROSITE; PS00584; PFKB_KINASES_2; 1.
DR PFAM; PF00294; pfkb; 1.
KW Transferase; Kinase.
FT CONFLICT 50 53 SMRV -> VDEG (IN REF. 2).
FT CONFLICT 62 69 LRALFGVR -> OGHSLVIE (IN REF. 2).
FT CONFLICT 108 108 Y -> D (IN REF. 2).
FT CONFLICT 290 314 LYMGAYTEENKROTHQLLPGAIA ->
FT CONFLICT 326 326 FIWRAIKRLNAKLSTHIYPRLL (IN REF. 2).
FT CONFLICT 326 326 R -> A (IN REF. 2).
SQ SEQUENCE 434 AA; 48479 MW; A0FF44EF CRC32;

Query Match
Best Local Similarity 73.0%; Score 46; DB 2; Length 434;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 379 PNSSKHVR 386
QY 119 PSSPKHV 126

RESULT 12
ID Q9WV44 PRELIMINARY; PRT; 531 AA.
AC Q9WV44;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE CYCLIN ANIA-6A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6;
RA BERKE J.D., HYMAN S.E.;
RT "A novel family of evolutionarily conserved cyclins potentially
RT involved in differentiation and neural plasticity.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF159159; AAD43568.1; -.
KW Cyclin.
SQ SEQUENCE 531 AA; 60564 MW; 458449D4 CRC32;

Query Match
Best Local Similarity 73.0%; Score 46; DB 11; Length 531;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 344 KPSSPREVK 352
QY 118 KPSSPKHVR 126

RESULT 13
ID Q9YX8 PRELIMINARY; PRT; 1175 AA.
AC Q9YX8;

```

US-09-376-430-2-15.rspt

Thu May 11 06:49:59 2000

```

DT 01-NOV-1999 (TREMELrel. 12, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMELrel. 12, Last annotation update)
DE REGULATOR OF G-PROTEIN SIGNALLING LOCO C2.
GN LOCO.

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RN YAMAZAKI J., KUSHIDA N., OGUCHI A., AOKI K., KUBOTA K., NAKAMURA Y.,
RX MEDLINE: 99180581.
RX KAWARABAYASHI Y., HINO Y., HORIKAWA H., YAMAZAKI S., HAIKAWA Y.,
RA JIN-NO K., TAKAHASHI M., SEKINE M., BABA S., ANKAI A., KOSUGI H.,
RA HOSOYAMA A., FUKUI S., NAGAI Y., NISHIJIMA K., NAKAZAWA H.,
RA TAKAMIYA M., MASUDA S., FUNAHASHI T., TANAKA T., KUDOH Y.,
RA YAMAZAKI J., KUSHIDA N., OGUCHI A., AOKI K., KUBOTA K., NAKAMURA Y.,
RA NOMURA N., SAKO Y., KIKUCHI H.;
RT "complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL: AP000059; BAA79432.1; -. 6AA7CEAE CRC32;
SQ SEQUENCE 144 AA; 15914 MW; 6AA7CEAE CRC32;

Query Match 73.0%; Score 46; DB 5; Length 1175;
Best Local Similarity 85.7%; Pred. No. 4.32e+00;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 61 STPKHVR 67
QY 120 SSPKHVR 126
|:|:|:|:|

RESULT 14
ID 001401 PRELIMINARY; PRT: 1861 AA.
AC 001401:
DT 01-JUL-1997 (TREMELrel. 04, Created)
DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)
DE MICROBUBULE ASSOCIATED PROTEIN.
GN ASP.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN-OREGON R;
RA SAUNDERS R.D.C., AVIDES M.C., HOWARD T.I.A., GONZALEZ C.,
RA GLOVER D.M.G.;
RL J. Cell Biol. 0:0-0(0).
DR EMBL: U95171; AAB51540.1; -.
DR FLYBASE: FBgn0000140; asp.
DR PFAM: PF00612; IQ: 12.
SQ SEQUENCE 1861 AA; 219558 MW; 12FDF4AF CRC32;

Query Match 73.0%; Score 46; DB 5; Length 1861;
Best Local Similarity 55.6%; Pred. No. 4.32e+00;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 471 KPATPKHVR 479
QY 118 KPSPKHVR 126
|:|:|:|:|

RESULT 15
ID Q9YEW4 PRELIMINARY; PRT: 144 AA.
AC Q9YEW4;
DT 01-NOV-1999 (TREMELrel. 12, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMELrel. 12, Last annotation update)
DE 144AA LONG HYPOTHETICAL PROTEIN.
GN APE0469.

```

```

OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Aeropyrum.
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN-K1;
RX MEDLINE: 99310339.
RX KAWARABAYASHI Y., HINO Y., HORIKAWA H., YAMAZAKI S., HAIKAWA Y.,
RA JIN-NO K., TAKAHASHI M., SEKINE M., BABA S., ANKAI A., KOSUGI H.,
RA HOSOYAMA A., FUKUI S., NAGAI Y., NISHIJIMA K., NAKAZAWA H.,
RA TAKAMIYA M., MASUDA S., FUNAHASHI T., TANAKA T., KUDOH Y.,
RA YAMAZAKI J., KUSHIDA N., OGUCHI A., AOKI K., KUBOTA K., NAKAMURA Y.,
RA NOMURA N., SAKO Y., KIKUCHI H.;
RT "complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL: AP000059; BAA79432.1; -. 6AA7CEAE CRC32;
SQ SEQUENCE 144 AA; 15914 MW; 6AA7CEAE CRC32;

Query Match 71.4%; Score 45; DB 1; Length 144;
Best Local Similarity 66.7%; Pred. No. 7.45e+00;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 17 KPASAKEVR 25
QY 118 KPSPKHVR 126
|:|:|:|:|

Search completed: Wed May 10 13:06:24 2000
Job time : 237 secs.

```





\*\*\*\*\*  
M P S R C H  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 12:33:51 2000; Maspar time 4.16 Seconds

Tabular output not generated.  
136.121 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (62-73) from US09376430A.pap (11 of 25)  
Perfect Score: 87  
Sequence: 1 FNGDEAYDQCTN 12

Scoring table:  
PAM 150  
Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: p1r62  
1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 23.985; Variance 31.123; scale 0.771

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	55	63.2	85	2	homeotic protein - co	1.82e+00
2	55	63.2	663	1	gelatinase A (EC 3.4.	1.82e+00
3	55	63.2	1380	2	ZMS1 protein - yeast	1.82e+00
4	52	59.8	2242	2	pyrimidine synthesis	7.14e+00
5	51	58.6	170	2	casein kinase II (EC	1.11e+01
6	51	58.6	348	2	casein kinase II (EC	1.11e+01
7	51	58.6	350	2	casein kinase II (EC	1.11e+01
8	51	58.6	350	2	casein kinase II (EC	1.11e+01
9	51	58.6	649	2	epithelial sodium cha	1.11e+01
10	51	58.6	650	2	sodium transport prot	1.11e+01
11	51	58.6	660	1	gelatinase A (EC 3.4.	1.11e+01
12	51	58.6	213	2	response regulator -	1.73e+01
13	50	57.5	350	2	casein kinase II (EC	1.73e+01
14	50	57.5	356	2	GTP-binding protein a	1.73e+01
15	50	57.5	360	2	casein kinase II (EC	1.73e+01
16	50	57.5	384	2	casein kinase II (EC	1.73e+01
17	50	57.5	391	2	casein kinase II (EC	1.73e+01
18	50	57.5	391	2	casein kinase (EC 2.7	1.73e+01
19	50	57.5	391	2	casein kinase II (EC	1.73e+01
20	50	57.5	391	2	casein kinase II (EC	1.73e+01
21	50	57.5	391	2	casein kinase II (EC	1.73e+01
22	50	57.5	398	2	aspartyl proteinase S	1.73e+01
23	50	57.5	551	1	NSL protein - vaccini	1.73e+01

24 50 57.5 551 2 A36848 N3L protein - variola 1.73e+01  
25 50 57.5 630 2 S50463 hypothetical protein 1.73e+01  
26 50 57.5 680 2 H70347 outer membrane protein 1.73e+01  
27 50 57.5 992 2 S49835 hypothetical protein 1.73e+01  
28 49 56.3 276 2 S41446 fibroin light chain ( 2.66e+01  
29 49 56.3 276 2 S41445 fibroin light chain ( 2.66e+01  
30 49 56.3 277 2 A33595 probable transposase 2.66e+01  
31 49 56.3 355 2 T10286 hypothetical protein 2.66e+01  
32 49 56.3 554 2 T15992 hypothetical protein 2.66e+01  
33 49 56.3 1356 2 T09107 DNA (cytosine-5'-) met 2.66e+01  
34 48 55.2 189 2 A43739 development-specific 4.07e+01  
35 48 55.2 205 2 A64671 ulcer-associated gene 4.07e+01  
36 48 55.2 394 2 F70197 hypothetical protein S1 4.07e+01  
37 48 55.2 411 2 A44121 ribosomal protein S1 4.07e+01  
38 48 55.2 437 2 T02689 hypothetical protein 4.07e+01  
39 48 55.2 561 2 S71189 Dwarfl protein - Arab 4.07e+01  
40 48 55.2 721 2 S37664 peplomeric polypeptide 4.07e+01  
41 48 55.2 700 2 A64925 Probable oxidoreducta 4.07e+01  
42 48 55.2 1102 2 S55100 hypothetical protein 4.07e+01  
43 48 55.2 1707 2 S77910 hypothetical protein 4.07e+01  
44 48 55.2 2109 1 I50421 aggreccan precursor - 4.07e+01  
45 48 55.2 4427 2 PN0637 polyketide synthase p 4.07e+01

## ALIGNMENTS

RESULT 1  
ENTRY T12636 #type fragment  
TITLE homeotic protein - common sunflower (fragment)  
ORGANISM #formal\_name Helianthus annuus #common\_name common sunflower  
DATE 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 13-Aug-1999  
ACCESSIONS T12636  
REFERENCE T17563  
#authors Chan, R.L.; Gonzalez, D.H.  
#journal Plant Physiol. (1994) 106:1687-1688  
#title A cDNA encoding an HD-zip protein from sunflower.  
#cross-references M01D:95148747  
#status preliminary; translated from GB/EMBL/DBJ  
#molecule\_type mRNA  
#residues 1-85 #label CHA  
#cross-references EMBL:L22848; NID:g349380; PID:g349381  
KEYWORDS DNA binding; homeobox; nucleus; transcription regulation  
SUMMARY #length 85 #checksum 7649

Query Match 63.2%; Score 55; DB 2; Length 85;  
Best Local Similarity 50.0%; Pred. No. 1.82e+00;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 26 YNSDEYENC 35  
QY 62 FNGDEAYDQC 71

RESULT 2

ENTRY S46492 #type complete  
TITLE gelatinase A (EC 3.4.24.24) precursor - chicken  
ORGANISM #formal\_name Gallus gallus #common\_name chicken  
DATE 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
ACCESSIONS S46492  
REFERENCE S46492  
#authors Aimes, R.T.; French, D.L.; Quigley, J.P.  
#journal Biochem. J. (1994) 300:729-736  
#title Cloning of a 72 kDa matrix metalloproteinase (gelatinase)  
from chicken embryo fibroblasts using gene family PCR:  
expression of the gelatinase increases upon malignant  
transformation.

#cross-references M01D:94280397  
#accession S46492  
#status preliminary  
#molecule\_type mRNA

```

##residues 1-663 ##label AIM
##cross-references EMBL:U07775; NID:g504475; PIDN:AAA19596.1;
          PID:g504476
##note
in the authors' translation 205-asp is shown after
residue 201 and, consequently, residues 202-204 are
displaced one codon to the right
CLASSIFICATION #superfamily gelatinase A; fibronectin type II repeat
homology; hemopexin repeat homology; matrix
metalloproteinase homology
KEYWORDS hydrolase; metalloproteinase; zinc; zymogen
FEATURE
67-216,391-443 #domain matrix metalloproteinase homology #status
atypical #label MMP\
230-271 #domain fibronectin type II repeat homology #label 2F1\
288-329 #domain fibronectin type II repeat homology #label 2F8\
346-387 #domain fibronectin type II repeat homology #label 2F9\
466-663 #domain hemopexin repeat homology #label PAX\
99,400,404,410 #binding_site zinc, catalytic (Cys, His, His, His)
(inhibited) #status predicted\
400,404,410 #binding_site zinc, catalytic (His) (active) #status
predicted\
401 #active_site Glu #status predicted
SUMMARY #length 663 #molecular-weight 74941 #checksum 837
Query Match 63.2%; Score 55; DB 1; Length 663;
Best Local Similarity 63.6%; Pred. No. 1.82e+00;
Matches 7; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
Db 294 FQGSYDQCT 303
| :|:|:|:|
QY 62 FNGDEAYDQCT 72

RESULT 3
ENTRY S57150 #type complete
TITLE ZMS1 protein - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES protein YJR127C
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 03-Sep-1995 #sequence_revision 01-Mar-1996 #text_change
06-Feb-1998
ACCESSIONS S57150; S43751
#authors Rose, M.; Koetter, P.; Entian, K.D.
#submission submitted to the Protein Sequence Database, September 1995
#accession S57150
#molecule_type DNA
##residues 1-1380 ##label ROS
##cross-references EMBL:249627; NID:g1015856; PID:g1015857; MIPS:YJR127C
REFERENCE S43751
#authors Thomas, D.; Barbey, R.; Surdin-Kerjan, Y.
#submission submitted to the EMBL Data Library, December 1993
#accession S43751
#molecule_type DNA
##residues 'MHTN'/18-1116,'IF',1119-1130,'H',1132-1142,'S' ##label
THO
##cross-references EMBL:L26506; NID:g432497; PID:g432498
GENETICS
#gene SGD:ZMS1
##cross-references SGD:S0003888; MIPS:YJR127C
#map_position.10R
SUMMARY #length 1380 #molecular-weight 155061 #checksum 1556
Query Match 63.28; Score 55; DB 2; Length 1380;
Best Local Similarity 66.74; Pred. No. 1.82e+00;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 1033 NGNEAYENC 1041
| :|:|:|
QY 63 NGDEAYDQC 71

RESULT 4.
ENTRY A57541 #type complete
##residues 1-170 ##label WAT
##experimental_source testis

TITLE pyrimidine synthesis multifunctional protein CAD - spiny
CONTAINS dogfish
aspartate carbamoyltransferase (EC 2.1.3.2);
carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC
6.3.5.5) large chain; dihydroorotase (EC 3.5.2.3)
#formal_name Squalus acanthias #common_name spiny dogfish
08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change
22-Jun-1999
ACCESSIONS A57541
REFERENCE A57541
#authors Hong, J.; Salo, W.L.; Anderson, P.M.
#journal J. Biol. Chem. (1995) 270:14130-14139
#title Nucleotide sequence and tissue-specific expression of the
multifunctional protein carbamoyl-phosphate
synthetase-aspartate transcarbamoylase-dihydroorotase (CAD)
mRNA in Squalus acanthias.
#cross-references MUID:95294021
#accession A57541
##status preliminary
##molecule_type mRNA
##residues 1-2242 ##label HON
##cross-references GB:U18868; NID:g951095; PIDN:AAA74569.1; PID:g951096
CLASSIFICATION #superfamily rudimentary enzyme; aspartate/ornithine
carbamoyltransferase homology; Bacillus dihydroorotase
homology; biotin carboxylase homology; carbamoyl-phosphate
synthase (ammonia) homology; carbamoyl-phosphate synthase
(glutamine-hydrolyzing) large chain homology;
carbamoyl-phosphate synthase (glutamine-hydrolyzing) small
chain homology; trpg homology
hydroxylase; ligase; transferase
#domain carbamoyl-phosphate synthase (ammonia) homology
#label CPA\
#domain carbamoyl-phosphate synthase
(glutamine-hydrolyzing) small chain homology #label
CPS\
178-354 #domain trpg homology #label TRG\
398-1446 #domain carbamoyl-phosphate synthase
(glutamine-hydrolyzing) large chain homology #label
CPL\
398-849 #domain biotin carboxylase homology #label BC1\
939-1385 #domain biotin carboxylase homology #label BC2\
1464-1808 #domain Bacillus dihydroorotase homology #label DHO\
1941-2239 #domain aspartate/ornithine carbamoyltransferase
homology #label ACT\
252 #active_site Cys #status predicted
SUMMARY #length 2242 #molecular-weight 249391 #checksum 857
Query Match 59.8%; Score 52; DB 2; Length 2242;
Best Local Similarity 40.0%; Pred. No. 7.14e+00;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Db 2169 FASDEEYAC 2178
| :|:|:|
QY 62 FNGDEAYDQC 71

RESULT 5
ENTRY PS0166 #type fragment
TITLE casein kinase II (EC 2.7.1.-) alpha' chain - bovine
(fragment)
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change
18-Jul-1997
ACCESSIONS PS0166
REFERENCE PS0166
#authors Watanabe, M.; Yuge, M.; Maeda, O.; Ohno, S.; Kawasaki, H.;
Suzuki, K.; Hidaka, H.
#submission submitted to JIPID, April 1991
#accession PS0166
#molecule_type mRNA
##residues 1-170 ##label WAT
##experimental_source testis

```

```

CLASSIFICATION #superfamily kinase-related transforming protein; protein
kinase homology
KEYWORDS ATP; heterotetramer; phosphotransferase;
serine/threonine-specific protein kinase
FEATURE
1-120 #domain protein kinase homology (fragment) #label KIN
SUMMARY
length 170 #checksum 3616
Query Match 58.6%; Score 51; DB 2; Length 170;
Best Local Similarity 55.6%; Pred. No. 1.11e+01;
Matches 3; Mismatches 1; Indels 0; Gaps 0;
Db 54 FHGQDNYDQ 62
QY 62 FNGDEAYDQ 70

RESULT 6
ENTRY casein kinase II (EC 2.7.1.1-) alpha' chain - zebra fish
TITLE #formal_name Brachydanio rerio #common_name zebra fish
ORGANISM 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change
DATE 18-Jun-1999
ACCESSIONS S74206; S78067
REFERENCE S74206
#authors Antonelli, M.; Daniotti, J.L.; Rojo, D.; Allende, C.C.;
Allende, J.E.
#journal Eur. J. Biochem. (1996) 241:272-279
#title Cloning, expression and properties of the alpha' subunit of
casein kinase 2 from zebrafish (Danio rerio).
#cross-references MUID:97054619
#accession S74206
#molecule_type mRNA
#residues 1-348 #label ANT
#cross-references EMBL:X99964
#experimental_source embryo
REFERENCE S78067
#authors Antonelli, M.; Daniotti, J.L.; Rojo, D.; Allende, C.C.;
Allende, J.E.
#submission submitted to the EMBL Data Library, August 1996
#description cloning, expression and properties of the alpha' subunit of
casein kinase CK2 from zebrafish (Danio rerio).
#accession S78067
#molecule_type mRNA
#residues 1-253,'Q',255-348 #label ANW
#cross-references EMBL:X99964; NID:G1495028; PIDN:CAA68229.1;
PID:G258484; PID:G1495029
CLASSIFICATION #superfamily kinase-related transforming protein; protein
kinase homology
KEYWORDS ATP; phosphotransferase; serine/threonine-specific protein
kinase
FEATURE
37-299 #domain protein kinase homology #label KIN\
45-53 #region protein kinase ATP-binding motif
SUMMARY
length 348 #molecular-weight 40862 #checksum 3627
Query Match 58.6%; Score 51; DB 2; Length 348;
Best Local Similarity 55.6%; Pred. No. 1.11e+01;
Matches 3; Mismatches 1; Indels 0; Gaps 0;
Db 233 FHGQDNYDQ 241
QY 62 FNGDEAYDQ 70

RESULT 7
ENTRY casein kinase II (EC 2.7.1.1-) alpha' chain - chicken
TITLE #formal_name Gallus gallus #common_name chicken
ORGANISM 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change
DATE 18-Jun-1999
ACCESSIONS B38611
REFERENCE A38611

```

```

#authors Maridor, G.; Park, W.; Krek, W.; Nigg, E.A.
#journal J. Biol. Chem. (1991) 266:2362-2368
#title Casein kinase II cDNA sequences, developmental expression,
and tissue distribution of mRNAs for alpha, alpha', and
beta subunits of the chicken enzyme.
#cross-references MUID:91115855
#accession B38611
#status preliminary
#molecule_type mRNA
#residues 1-350 #label MAR
#cross-references GB:M59457; GB:J05737; NID:G211521; PIDN:AAA48686.1;
PID:G211522
CLASSIFICATION #superfamily kinase-related transforming protein; protein
kinase homology
KEYWORDS ATP; heterotetramer; phosphotransferase;
serine/threonine-specific protein kinase
FEATURE
38-300 #domain protein kinase homology #label KIN\
46-54 #region protein kinase ATP-binding motif
SUMMARY
length 350 #molecular-weight 41246 #checksum 4375
Query Match 58.6%; Score 51; DB 2; Length 350;
Best Local Similarity 55.6%; Pred. No. 1.11e+01;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 234 FHGQDNYDQ 242
QY 62 FNGDEAYDQ 70

RESULT 8
ENTRY casein kinase II (EC 2.7.1.1-) alpha' chain - human
TITLE #formal_name Homo sapiens #common_name man
ORGANISM 31-Oct-1990 #sequence_revision 13-Jan-1993 #text_change
DATE 18-Jun-1999
ACCESSIONS B35838
REFERENCE A35838
#authors Lozeman, F.J.; Litchfield, D.W.; Piensing, C.; Taklo, K.;
Walsh, K.A.; Krebs, E.G.
#journal Biochemistry (1990) 29:8436-8447
#title Isolation and characterization of human cDNA clones encoding
the alpha and the alpha' subunits of casein kinase II.
#cross-references MUID:91070071
#accession B35838
#status preliminary
#molecule_type mRNA
#residues 1-350 #label LOZ
#cross-references GB:M55288; GB:J02924; NID:G177837; PIDN:AAA51548.1;
PID:G177838
#note the authors translated the codon CAA for residue 37 as
Asn, GAG for residue 64 as Ile, AAG for residue 199 as
Leu, and CAC for residue residue 263 as Gly
GENETICS
#gene GDB:CSNK2A2; CSNK2A1
#cross-references GDB:129561; OMIM:115442
#map_position 16q13-16q13
CLASSIFICATION #superfamily kinase-related transforming protein; protein
kinase homology
KEYWORDS ATP; heterotetramer; phosphotransferase;
serine/threonine-specific protein kinase
FEATURE
38-300 #domain protein kinase homology #label KIN\
46-54 #region protein kinase ATP-binding motif
SUMMARY
length 350 #molecular-weight 41213 #checksum 1991
Query Match 58.6%; Score 51; DB 2; Length 350;
Best Local Similarity 55.6%; Pred. No. 1.11e+01;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 234 FHGQDNYDQ 242
QY 62 FNGDEAYDQ 70

```

```

RESULT 9
ENTRY I64847 #type complete
TITLE epithelial sodium channel gamma subunit - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 29-May-1998 #sequence_revision 29-May-1998 #text_change
24-Sep-1999
ACCESSIONS I64847
REFERENCE I51915
#authors McDonald, F.J.; Snyder, P.M.; Price, M.P.; Welsh, M.J.
#journal Am. J. Physiol. (1995) 268:1157-1163
#title Cloning and expression of the beta and gamma subunits of the
#accession human epithelial sodium channel.
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-649 #label RES
#cross-references GB:L36592; NID:g845511; PID:g845512
GENETICS gamma hENaC
#gene
CLASSIFICATION #superfamily fibronectin type I repeat homology
SUMMARY #length 649 #molecular_weight 74342 #checksum 3362
Query Match 58.6%; Score 51; DB 2; Length 649;
Best Local Similarity 50.0%; Pred. No. 1.11e+01;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Db 363 FKLSEPYSCQE 374
I: 1:1111:
QY 62 FNGDEAYDOCTN 73

RESULT 10
ENTRY A54065 #type complete
TITLE sodium transport protein gamma chain - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change
24-Sep-1999
ACCESSIONS A54065; S41160
REFERENCE A54065
#authors Lingueglia, E.; Renard, S.; Waldmann, R.; Voilley, N.;
#journal Champigny, G.; Plass, H.; Lazdunski, M.; Barbry, P.
#title J. Biol. Chem. (1994) 269:13736-13739
#accession Different homologous subunits of the amiloride-sensitive Na
#cross-references MUID:94245676
#status preliminary
#molecule_type mRNA
#residues 1-650 #label LIN
#cross-references GB:X78034; NID:g495270; PID:g495271
REFERENCE S41158
#authors Canessa, C.M.; Schild, L.; Buell, G.; Thorens, B.; Gautschi,
#journal I.; Horisberger, J.D.; Rossier, B.C.
#title Nature (1994) 367:463-467
#accession Amiloride-sensitive epithelial Na(+) channel is made of three
#cross-references MUID:94150824
#status preliminary
#molecule_type mRNA
#residues 1-52, P' 54-572, C' 574-650 #label CAN
#cross-references EMBL:X77933; NID:g458849; PID:g458850
CLASSIFICATION #superfamily fibronectin type I repeat homology
SUMMARY #length 650 #molecular_weight 74065 #checksum 7454
Query Match 58.6%; Score 51; DB 2; Length 650;
Best Local Similarity 50.0%; Pred. No. 1.11e+01;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Db 364 FKLSEPYSCQE 375

```

```

QY 62 FNGDEAYDOCTN 73
I: 1:1111:

RESULT 11
ENTRY A28153 #type complete
TITLE gelatinase A (EC 3.4.24.24) precursor - human
ALTERNATE_NAMES collagenase type IV; matrix metalloproteinase 2 (MMP2);
progelatinase A
ORGANISM #formal_name Homo sapiens #common_name man
DATE 28-Aug-1989 #sequence_revision 07-Jul-1995 #text_change
18-Jun-1999
ACCESSIONS A28153; A34202; A42225; A60187; S13858; S39436; A31480;
S44432; A61498; S55327; S13953
REFERENCE A28153
#authors Collier, I.E.; Wilhelm, S.M.; Eisen, A.Z.; Marmer, B.L.;
Grant, G.A.; Seltzer, J.L.; Kronberger, A.; He, C.; Bauer,
E.A.; Goldberg, G.I.
#journal J. Biol. Chem. (1988) 263:6579-6587
#title H-ras oncogene-transformed human bronchial epithelial cells
#accession (FBE-1) secrete a single metalloprotease capable of
degrading basement membrane collagen.
#cross-references MUID:88198218
#accession A28153
#molecule_type mRNA
#residues 30-660 #label COL
#cross-references GB:J03210; NID:g180670; PIDN:AAA35701.1; PID:g180671
REFERENCE A34202
#authors Huhtala, P.; Eddy, R.L.; Fan, Y.S.; Byers, M.G.; Shows, T.B.;
Tryggvason, K.
#journal Genomics (1990) 6:554-559
#title Completion of the primary structure of the human type IV
collagenase preproenzyme and assignment of the gene (CLG4)
to the q21 region of chromosome 16.
#cross-references MUID:90228972
#accession A34202
#molecule_type DNA
#residues 1-51 #label HU2
#cross-references GB:M33789; NID:g180600; PIDN:AAA52027.1; PID:g180601
REFERENCE A42225
#authors Huhtala, P.; Chow, L.T.; Tryggvason, K.
#journal J. Biol. Chem. (1990) 265:11077-11082
#title Structure of the human type IV collagenase gene.
#cross-references MUID:90293047
#accession A42225
#status not compared with conceptual translation
#molecule_type DNA
#residues 1-51;220-393 #label HUH
#cross-references GB:M5593; GB:J05471; NID:g180614; PIDN:AAA52028.1;
PID:g180616
#note neither the complete amino acid nor the complete
nucleotide sequence is given in this paper
REFERENCE A60187
#authors Frisch, S.M.; Reich, R.; Collier, I.E.; Genrich, L.T.;
Martin, G.; Goldberg, G.I.
#journal Oncogene (1990) 5:75-83
#title Adenovirus E1A represses protease gene expression and
inhibits metastasis of human tumor cells.
#cross-references MUID:90206614
#accession A60187
#status not compared with conceptual translation
#molecule_type DNA
#residues 1-58 #label FRI
REFERENCE S13858
#authors Okada, Y.; Morodomi, T.; Enghild, J.J.; Suzuki, K.; Yasui,
A.; Nakanishi, I.; Salvesen, G.; Nagase, H.
#journal Eur. J. Biochem. (1990) 194:721-730
#title Matrix metalloproteinase 2 from human rheumatoid synovial
fibroblasts. Purification and activation of the precursor
and enzymic properties.
#cross-references MUID:91099351
#accession S13858
#molecule_type protein

```

US-09-376-430-2-11.rpr

Thu May 11 06:49:49 2000

```

#residues 30-45:110-124 #label OKA
REFERENCE S39436
#authors Crabbe, T.; Ioannou, C.; Docherty, A.J.P.
#journal Eur. J. Biochem. (1993) 218:431-438
#title Human progelatinase A can be activated by autolysis at a rate that is concentration-dependent and enhanced by heparin bound to the C-terminal domain.
#cross-references MUID:94094834
#accession S39436
#molecule_type protein
#residues 30-44:444-456 #label CR2
REFERENCE A31480
#authors Stetler-Stevenson, W.G.; Kruttsch, H.C.; Wachter, M.P.; Margulies, I.M.K.; Liotta, L.A.
#journal J. Biol. Chem. (1989) 264:1353-1356
#title The activation of human type IV collagenase proenzyme. Sequence identification of the major conversion product following organomercurial activation.
#cross-references MUID:89109136
#accession A31480
#molecule_type protein
#residues 110-123 #label STE
REFERENCE S44432
#authors Crabbe, T.; Smith, B.; O'Connell, J.; Docherty, A.
#journal FEBS Lett. (1994) 345:14-16
#title Human progelatinase A can be activated by matrixlysin.
#cross-references MUID:94252395
#accession S44432
#molecule_type protein
#residues 110-115 #label CRA
REFERENCE A61496
#authors Brown, D.; Chwa, M.; Escobar, M.; Kenney, M.C.
#journal Exp. Eye Res. (1991) 52:5-16
#title Characterization of the major matrix degrading metalloproteinase of human corneal stroma. Evidence for an enzyme/inhibitor complex.
#cross-references MUID:91330998
#accession A61496
#molecule_type protein
#residues 'X', '31', 'X', '33-46', 'X', '48-50', 'Q' #label BRO
#experimental_source corneal stroma
REFERENCE S53327
#authors Itoh, Y.; Blinner, S.; Nagase, H.
#journal Biochem. J. (1995) 308:645-651
#title Steps involved in activation of the complex of pro-matrix metalloproteinase 2 (progelatinase A) and tissue inhibitor of metalloproteinases (TIMP)-2 by 4-aminophenylmercuric acetate.
#cross-references MUID:95290003
#accession S53327
#molecule_type protein
#residues 110-114 #label ITO
GENETICS
#gene GDB:MMP2; CLG4; CLG4A
#map_position 16q13-16q13
#map_position 51/3; 127/2; 178/1; 220/1; 278/1; 336/1; 394/1; 446/1; 491/2; 537/1; 590/2; 627/1
#introns
FUNCTION
#description proteolytic cleavage of gelatin type I and collagen types IV, VII, and X
CLASSIFICATION
#superfamily gelatinase A; fibronectin type II repeat. homology; hemopexin repeat homology; matrix metalloproteinase homology
KEYWORDS
#domain activation peptide #status predicted #label ACT
#domain matrix metalloproteinase homology #status atypical #label MMP
#product gelatinase A #status predicted #label MAT
FEATURE
1-29 #domain signal sequence #status predicted #label SIG
30-660 #product progelatinase A #status predicted #label PRO
70-109 #domain activation peptide #status predicted #label ACT
70-219, 394-446 #domain matrix metalloproteinase homology #status atypical #label MMP
110-660 #product gelatinase A #status predicted #label MAT

```

```

233-390 #region collagen binding #status predicted\
233-274 #domain fibronectin type II repeat homology #label 2F1\
291-332 #domain fibronectin type II repeat homology #label 2F8\
349-390 #domain fibronectin type II repeat homology #label 2F9\
463-660 #domain hemopexin repeat homology #label PXN\
102,403,407,413 (binding site zinc, catalytic (Cys, His, His) (inhibited) #status predicted\
403,407,413 (binding site zinc, catalytic (His) (active) #status predicted\
404 #active site Glu #status predicted\
469-660 #disulfide bonds #status predicted\
573,642 #binding site carbohydrate (Asn) (covalent) #status predicted\
SUMMARY #length 660 #molecular-weight 73882 #checksum 9220
Query Match 58.6%; Score 51; DB 1; Length 660;
Best Local Similarity 58.3%; Pred. No. 1,11e+01;
Matches 7; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
Db 239 FNGKE-YNSCTD 249
QY 62 FNGDEAYDOCTN 73
RESULT 12
ENTRY #type complete
TITLE response regulator - Helicobacter pylori (strain 26695)
ORGANISM #formal_name Helicobacter pylori
DATE 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 26-Aug-1999
ACCESSION E64690
REFERENCE A64520
#authors Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Ketchum, K.A.; Klenk, H.P.; Gill, S.; Dougherty, B.A.; Nelson, K.; Quackenbush, J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, K.; Fitzgerald, L.M.; Lee, N.; Adams, M.D.; Hickey, E.K.; Berg, D.E.; Gocayne, J.D.; Utterback, T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.M.; Venter, J.C.
#journal Nature (1997) 388:539-547
#title The complete genome sequence of the gastric pathogen Helicobacter pylori.
#cross-references MUID:97394467
#accession E64690
#status preliminary; nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-213 #label TOM
#cross-references GB:AE000636; GB:AE000511; NID:G2314517; PID:G2314531; TIGR:HP1365
CLASSIFICATION #superfamily ompR protein; response regulator homology
KEYWORDS phosphoprotein
FEATURE
5-115 #domain response regulator homology #label RRH\
53 #binding site phosphate (Asp) (covalent) #status predicted\
SUMMARY #length 213 #molecular-weight 24871 #checksum 3023
Query Match 57.5%; Score 50; DB 2; Length 213;
Best Local Similarity 66.7%; Pred. No. 1.7e+01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 33 FNGKEAYER 41
QY 62 FNGDEAYDO 70
RESULT 13
ENTRY #type complete
S20404

```

```

TITLE      casein kinase II (EC 2.7.1.1-) alpha chain - African clawed
ORGANISM    frog
DATE        #formal_name Xenopus laevis #common_name African clawed frog
            #sequence_revision 19-Oct-1995 #text_change
            18-Jun-1999
ACCESSIONS  S20404; S18897
REFERENCE    S20404; S18897
#authors    Jedlicki, A.; Hinrichs, M.V.; Allende, C.C.; Allende, J.E.
#journal     FEBS Lett. (1992) 297:280-284
#title       The cDNAs coding for the alpha- and beta-subunits of Xenopus
            laevis casein kinase II.
#cross-references MUID:92183811
#accession   S20404
#status      preliminary
#molecule_type mRNA
#residues    1-350 #label JED
#cross-references EMBL:X62375; NID:g64627; PIDN:CAA44238.1; PID:g64628
CLASSIFICATION #superfamily kinase-related transforming protein; protein
            kinase homology
KEYWORDS     ATP; heterotetramer; phosphotransferase;
            serine/threonine-specific protein kinase
FEATURE      37-299      #domain protein kinase homology #label KIN
            45-53      #region protein kinase ATP-binding motif
SUMMARY      #length 350 #molecular-weight 41454 #checksum 1690
            Query Match 57.5%; Score 50; DB 2; Length 350;
            Best Local Similarity 55.6%; Pred. No. 1.73e+01;
            Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
            Db 233 FHGHDNYDQ 241
            QY 62 FNGDEAYDQ 70
            I:I : I:I
            I:I : I:I

RESULT 14
ENTRY   A32945      #type complete
TITLE   GTP-binding protein alpha-1 chain - slime mold (Dictyostelium
            discoideum)
ORGANISM #formal_name Dictyostelium discoideum
DATE     20-Dec-1989 #sequence_revision 20-Dec-1989 #text_change
            13-Aug-1999
ACCESSIONS A32945
REFERENCE   A32945
#authors    Pupillo, M.; Kumagai, A.; Pitt, G.S.; Firtel, R.A.;
            Devreotes, P.N.
#journal     Proc. Natl. Acad. Sci. U.S.A. (1989) 86:4892-4896
#title       Multiple alpha-subunits of guanine nucleotide-binding
            proteins in Dictyostelium.
#cross-references MUID:89296910
#accession   A32945
#status      preliminary
#molecule_type mRNA
#residues    1-356 #label PUP
#cross-references GB:M25060; NID:gi67792; PIDN:AAA33207.1; PID:gi67793
CLASSIFICATION #superfamily GTP-binding regulatory protein Gs alpha chain
            GTP binding; P-loop
KEYWORDS     43-50      #region nucleotide-binding motif A (P-loop)\
            271-274      #region GTP-binding NKXD motif
SUMMARY      #length 356 #molecular-weight 40187 #checksum 678
            Query Match 57.5%; Score 50; DB 2; Length 356;
            Best Local Similarity 25.0%; Pred. No. 1.73e+01;
            Matches 3; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
            Db 292 YDGPQTYEGCSE 303
            QY 62 FNGDEAYDQCTN 73
            :I: :I: :I:
            :I: :I: :I:

RESULT 15
ENTRY   A35562      #type complete

```

```

TITLE      casein kinase II (EC 2.7.1.1-) alpha chain - Caenorhabditis
ORGANISM    elegans
DATE        #formal_name Caenorhabditis elegans
            31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change
            18-Jun-1999
ACCESSIONS  A35562
REFERENCE    A35562
#authors    Hu, E.; Rubin, C.S.
#journal     J. Biol. Chem. (1990) 265:5072-5080
#title       Casein kinase II from Caenorhabditis elegans. Properties and
            developmental regulation of the enzyme; cloning and
            sequence analyses of cDNA and the gene for the catalytic
            subunit.
#cross-references MUID:90202988
#accession   A35562
#molecule_type mRNA
#residues    1-360 #label HUX
#cross-references GB:J05274; NID:gi56247; PIDN:AAA27984.1; PID:gi56248
CLASSIFICATION #superfamily kinase-related transforming protein; protein
            kinase homology
KEYWORDS     ATP; heterotetramer; phosphotransferase;
            serine/threonine-specific protein kinase
FEATURE      36-298      #domain protein kinase homology #label KIN\
            44-52      #region protein kinase ATP-binding motif
SUMMARY      #length 360 #molecular-weight 42257 #checksum 5512
            Query Match 57.5%; Score 50; DB 2; Length 360;
            Best Local Similarity 55.6%; Pred. No. 1.73e+01;
            Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
            Db 232 FHGHDNYDQ 240
            QY 62 FNGDEAYDQ 70
            I:I : I:I
            I:I : I:I

Search completed: Wed May 10 12:33:59 2000
Job time : 8 secs.

```

\*\*\*\*\*  
 M P S R L H (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 13:15:26 2000; MasPar time 2.84 Seconds  
 Tabular output not generated. 66,637 Million cell updates/sec

Title: >US-09-376-430-2  
 Description: (139-146) from US09376430A.ppt (16 of 25)  
 Perfect score: 55  
 Sequence: 1 SDSLSDGL 8

Scoring table: PAM 150  
 Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: a-geneseq35  
 1:geneseqp

Statistics: Mean 16.021; Variance 45.326; scale 0.353  
 Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Description	ID	Pred. No.
1	47	85.5	ACVS.	R13753	1.12e+02
2	47	85.5	ACV synthetase.	R13895	1.12e+02
3	45	81.8	Bovine poly-immunoglob	W03179	1.88e+02
4	44	80.0	Human cytomegalovirus	W27083	2.43e+02
5	44	80.0	Human UL105 open readin	W27086	2.43e+02
6	44	80.0	Human IRS-1 and IRS-2	W93972	2.43e+02
7	44	80.0	Insulin receptor subst	R67708	2.43e+02
8	42	76.4	Bovine RFBKP.	R28979	4.03e+02
9	42	76.4	Human FKBP-13 immunoph	R93551	4.03e+02
10	42	76.4	HRFBKP.	R28980	4.03e+02
11	42	76.4	H. pylori cytoplasmic	W24593	4.03e+02
12	42	76.4	H. pylori cytoplasmic	W20146	4.03e+02
13	42	76.4	H. pylori ORF 05p1081	W10982	4.03e+02
14	42	76.4	H. pylori cytoplasmic	W20740	4.03e+02
15	42	76.4	Protein encoded by ORF	W21783	4.03e+02
16	42	76.4	Herpesvirus of turkeys	W03136	4.03e+02
17	42	76.4	Herpesvirus of turkeys	W03332	4.03e+02
18	42	76.4	L.lactis branched amin	R54216	4.03e+02
19	42	76.4	ShE2 enterotoxin enco	R76593	4.03e+02
20	42	76.4	Eier enterotoxin enco	R76700	4.03e+02
21	42	76.4	T. harzianum exochitin	W01639	4.03e+02
22	42	76.4	Hexosaminidase enzym	W85604	4.03e+02
23	42	76.4	Helicobacter polypepti	W71477	4.03e+02

24	42	76.4	900	1	R20568	Sequence of a partial	4.03e+02
25	42	76.4	925	1	W55729	H. pylori ORF 05p1081	4.03e+02
26	42	76.4	966	1	W19916	Drosophila melanogaste	4.03e+02
27	42	76.4	972	1	W34497	Obesity receptor A pro	4.03e+02
28	42	76.4	973	1	W48304	Amino acid sequence of	4.03e+02
29	42	76.4	1003	1	W19917	Drosophila virilis Ksr	4.03e+02
30	41	74.5	102	1	W41157	Biological clock contr	5.17e+02
31	41	74.5	228	1	R88275	Papilloma virus major	5.17e+02
32	41	74.5	260	1	R28871	Odorant receptor clone	5.17e+02
33	41	74.5	269	1	R48743	G-protein coupled odor	5.17e+02
34	41	74.5	274	1	R48748	G-protein coupled odor	5.17e+02
35	41	74.5	275	1	W02717	G-protein coupled odor	5.17e+02
36	41	74.5	277	1	W02719	G-protein coupled odor	5.17e+02
37	41	74.5	284	1	W02718	G-protein coupled odor	5.17e+02
38	41	74.5	312	1	R27875	Odorant receptor clone	5.17e+02
39	41	74.5	314	1	R27874	Odorant receptor clone	5.17e+02
40	41	74.5	464	1	W88311	E. coli O111 antigen g	5.17e+02
41	41	74.5	513	1	Y03186	Rat Acid sensitive ion	5.17e+02
42	41	74.5	559	1	W68507	Rat Acid sensing ionic	5.17e+02
43	41	74.5	566	1	W20945	H. pylori inner membra	5.17e+02
44	41	74.5	599	1	R75655	Human afamin.	5.17e+02
45	41	74.5	879	1	W72979	Bovine beta-mannosidas	5.17e+02

ALIGNMENTS

RESULT 1  
 ID R13753 standard; Protein: 3768 AA.  
 AC R13753.  
 DT 14-NOV-1991 (first entry)  
 DE ACVS.  
 KW Delta-(L-alpha-aminoadipyl)-L-cysteiny-D-valine synthetase;  
 OS Beta-lactam antibiotics.  
 OS Penicillium chrysogenum.  
 FH Key Location/Qualifiers  
 FT domain 301..1068  
 FT domain /label= domain I  
 FT domain 1392..2154  
 FT domain /label= domain II  
 FT domain 2474..3295  
 FT domain /label= domain III  
 PN EP-444758-A.  
 PD 04-SEP-1991.  
 PF 27-FEB-1991; 200422.  
 PR 28-FEB-1990; EP-200475.  
 PR 28-FEB-1990; EP-200488.  
 PR 02-JUL-1990; EP-201768.  
 PR 03-OCT-1990; EP-202628.  
 PR 27-FEB-1991; EP-200422.  
 PA (KONN ) GIST-BROCADES NV.  
 PI Veenstra AE, Martin JF, Garcia BD, Gutierrez S, Barredo JL,  
 PI Montenegro Prieto E, Von Doehren H, Palissa H, Van Liempt H;  
 PI WPI; 91-263525/36.  
 DR N-PDSB; Q13547.  
 DR Mutant delta-(L-alpha-aminoadipyl)-L-cysteiny-D-valine  
 PT synthetase - used in prodn. of beta-lactam antibiotics.  
 PS Disclosure; Page 19; 56pp; English.  
 CC The amino acid sequence codes for delta- (L-alpha-aminoadipyl)-L-  
 CC cysteinyl-D-valine-synthetase (ACVS). The prods. may be used for  
 CC the enhanced expression (in vivo and in vitro) of mutant enzymes  
 CC and fermentable or known and new beta- lactam antibiotics and their  
 CC precursors, partic. antibiotics of the penam and cephem classes.  
 CC See also R13754-R13756.  
 SQ Sequence 3768 AA;

Query Match: 85.5% Score 47; DB 1; Length 3768;  
 Best Local Similarity 57.1%; Pred. No. 1.12e+02;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 337 ELTYGEL 343  
 QY 140 DLSYDGL 146

RESULT 2  
 ID R13895 standard; Protein; 3778 AA.  
 AC R13895;  
 DT 22-NOV-1991 (first entry)  
 DE ACV synthetase.  
 KW Beta lactam antibiotics; penicillin.  
 OS Penicillin chrysoygenum.  
 FH Key  
 FT domain  
 FT 301..1068  
 FT Location/Qualifiers  
 FT /label= I  
 FT /function= activation of amino acid substrate  
 FT 374..423  
 FT /label= subdomain  
 FT 474..501  
 FT /label= subdomain  
 FT 655..699  
 FT /label= subdomain  
 FT 725..754  
 FT /label= subdomain  
 FT 1392..2154  
 FT /label= II  
 FT /function= activation of amino acid substrate  
 FT 1470..1518  
 FT /label= subdomain  
 FT 1564..1590  
 FT /label= subdomain  
 FT 1745..1789  
 FT /label= subdomain  
 FT 1817..1846  
 FT /label= subdomain  
 FT 2474..3295  
 FT /label= III  
 FT /function= activation of amino acid substrate  
 FT 2554..2603  
 FT /label= subdomain  
 FT 2647..2673  
 FT /label= subdomain  
 FT 2827..2871  
 FT /label= subdomain  
 FT 2899..2928  
 FT /label= subdomain  
 FT 3560..3647  
 FT /label= IV  
 FT /function= thioesterase  
 PN EP-445868-A.  
 PD 11-SEP-1991.  
 PF 27-FEB-1991; 200423.  
 PR 28-FEB-1990; EP-200475.  
 PR 28-FEB-1990; EP-200488.  
 PR 02-JUL-1990; EP-201768.  
 PR 03-OCT-1990; EP-202628.  
 PR 27-FEB-1991; EP-200423.  
 PR (KONN ) GIST-BROCADES NV.  
 PA Veenstra AE, Martin JF, Garcia BD, Guttierrez S, Barredo JL;  
 PI Montenegro PE, Von Doehren H, Palissa H, Van Liempt H;  
 DR WPI; 91-268755/37.  
 DR N-PSDB; Q13607.  
 PT DNA encoding amino:adipyl-cysteinyI-valine synthetase - used for  
 PT prodn. of the enzyme or enhanced prodn. of new or known  
 PT beta-lactam antibiotic cpds.  
 PS Claim 1; Page 20; 54pp; English.  
 CC The isolation of DNA fragments encoding the ACV synthetase is  
 CC described in EP-357119. Cosmid HM193 contains one such fragment  
 CC which was sequenced using the Sequenase system 2.0. The protein  
 CC sequence was deduced from the DNA. Three distinct regions of  
 CC homology have been identified, domains I, II and III. Within  
 CC these domains several even more conserved elements can be  
 CC distinguished. Since the enzyme synthesises a tripeptide, which  
 CC most probably requires the activation of three amino acids, a  
 CC role for these domains in the activation reactions seems likely.  
 CC A fourth domain is thought to act as a thioesterase.  
 CC The gene can be used to express the synthetase enzyme which can

CC be used for the prodn of new beta-lactam antibiotics.  
 CC See also R13896.  
 SQ Sequence 3778 AA;

Query Match 85.5%; Score 47; DB 1; Length 3778;  
 Best Local Similarity 57.1%; Pred. No. 1.12e+02;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 347 ELTYGEL 353  
 Qy 140 DLSIGDL 146

## RESULT 3

ID W03179 standard; Protein; 757 AA.  
 AC W03179;  
 DT 24-FEB-1997 (first entry)  
 DE Bovine poly-immunoglobulin receptor.  
 KW Bovine; immunoglobulin; receptor; protection protein; mutants;  
 KW heavy chain; antigen binding domain; protection; pathogen;  
 KW mucosal; environment; gastrointestinal; passive; immunisation;  
 KW Guy's 13 antibody; prevention; dental caries; Streptococcus;  
 KW poly; sorbinus; cow.  
 OS Bos taurus.  
 FH Key  
 FT region  
 FT 13..45  
 FT Location/Qualifiers  
 FT /note= "putative immunoglobulin binding residues  
 FT of domain I"  
 FT 1..120  
 FT /label= domain\_I  
 FT 110..230  
 FT /label= domain\_II  
 FT 210..340  
 FT /label= domain\_III  
 FT 320..450  
 FT /label= domain\_IV  
 FT 440..550  
 FT /label= domain\_V  
 FT 550..606  
 FT /note= "external portions of domain VI"  
 FT 550..627  
 FT /note= "external portions of domain VI"  
 FT 625..660  
 FT /label= transmembrane\_segment  
 FT 650..757  
 FT /label= intracellular\_portion  
 PN W09621012-A1.  
 PD 11-JUL-1996.  
 PF 27-DEC-1995; U16889.  
 PR 30-DEC-1994; US-367395.  
 PR 04-MAY-1995; US-434000.  
 PA (PLAN-) PLANT BIOTECHNOLOGY INC.  
 PA (UNNE-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.  
 PA (PLAN-) PLANET BIOTECHNOLOGY INC.  
 PI Hiatt AC, Lehner T, Ma JKC;  
 DR WPI; 96-333987/33.  
 DR N-PSDB; T31289.  
 PT Immunoglobulin and protection protein complex and its prodn. in  
 PT plants - useful for passive immunisation against mucosal antigens,  
 PT esp. against S. mutans and S. sorbinus to prevent dental caries  
 PT Disclosure; Pages 111-115; 152pp; English.  
 CC The present sequence is the bovine poly-immunoglobulin (Ig)  
 CC receptor, a portion of which corresp. to residues 1-627, pref.  
 CC 1-606, or esp. residues 13-45, 1-120, 110-230, 210-340, 320-450,  
 CC 440-550, 550-606 or 550-627 comprises a protection protein (PP).  
 CC The Ig of the invention comprises a PP as above in association with  
 CC an Ig derived heavy chain, having at least a portion of an antigen  
 CC (Ag) binding domain. The PP protects the Ig in harsh mucosal, e.g.  
 CC gastrointestinal, environments, therefore enhancing its  
 CC effectiveness in passively immunising animals against mucosal  
 CC pathogens. The Ag binding domain is specifically derived from the  
 CC Guy's 13 antibody, and the Ig can be used to prevent dental caries  
 CC by binding, e.g. Streptococcus mutans serotypes c, e and f, or



CC S. sorbinus serotypes d and g.  
 SQ Sequence 757 AA;  
 Query Match 81.8%; Score 45; DB 1; Length 757;  
 Best Local Similarity 62.5%; Pred. No. 1.88e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 242 PELVYGD 249  
 QY 139 SLSYGD 146  
 RESULT 4  
 ID W27083 standard; Protein; 956 AA.  
 AC W27083;  
 DT 18-MAR-1998 (first entry)  
 DE Human cytomegalovirus helicase.  
 KW Human cytomegalovirus helicase; HCMV UL105; open reading frame; orf;  
 screening; inhibitor; infection; diagnosis; antiviral.  
 OS Homo sapiens.  
 PN GB2311068-A.  
 PD 17-SEP-1997.  
 PF 05-MAR-1997; 004575.  
 PR 04-APR-1996; GB-007118.  
 PR 14-MAR-1996; US-013389.  
 PA (MERI) MERCK & CO INC.  
 PI Gotlib L, Hazuda DJ, Lafemina RL;  
 DR WPI; 97-427907/40.  
 DR P-PSDB; W27083.  
 PT Novel human cytomegalovirus helicase - used for screening for HCMV  
 antivirals and in diagnosis of HCMV related diseases  
 PS Claim 1; Pages 13-16; 28pp; English.  
 CC The present sequence represents a new human cytomegalovirus (HCMV)  
 helicase which was used in a screening assay for compounds which  
 inhibit HCMV helicase, preferably with an IC50 of not greater than 200nM.  
 CC DNA encoding HCMV helicase, spanning nucleotides 151926 through 154793 of  
 HCMV AD169 was amplified by PCR primers T85063-4. The amplified product  
 CC was digested with BglII and EcoRI and cloned into BglII/EcoRI digested  
 CC pB877 vector DNA to yield the plasmid pB877 UL105. The amplified HCMV  
 CC UL105 orf sequence and its translational product, helicase, were  
 CC determined. The HCMV UL105 orf was recombined into the baculovirus AcNPV  
 CC genome by standard co-transfection protocols. This new HCMV helicase  
 CC is also used as a diagnostic tool for diseases resulting from HCMV  
 CC infection. 956 AA;  
 SQ Sequence 956 AA;  
 Query Match 80.0%; Score 44; DB 1; Length 956;  
 Best Local Similarity 62.5%; Pred. No. 2.43e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 359 TDLDFGDL 366  
 QY 139 SLSYGD 146  
 RESULT 5  
 ID W27086 standard; Protein; 956 AA.  
 AC W27086;  
 DT 18-MAR-1998 (first entry)  
 DE HCMV UL105 open reading frame translation product.  
 KW Human cytomegalovirus primase; HCMV UL105; open reading frame; orf;  
 screening; inhibitor; infection; diagnosis; ss.  
 OS Homo sapiens.  
 PN GB2311069-A.  
 PD 17-SEP-1997.  
 PF 05-MAR-1997; 004577.  
 PR 04-APR-1996; GB-007117.  
 PR 14-MAR-1996; US-013546.  
 PA (MERI) MERCK & CO INC.  
 PI Gotlib L, Hazuda DJ, Lafemina RL;  
 DR WPI; 97-427907/40.  
 DR N-PSDB; T85076.  
 PT Novel human cytomegalovirus primase - used for screening for HCMV

PT antivirals and in diagnosis of HCMV related diseases  
 PS Example 1; Pages 17-20; 28pp; English.  
 CC This sequence represents the translation product of human cytomegalovirus  
 (HCMV) UL105 open reading frame (orf) spanning nucleotides 151926 through  
 CC 154793 of HCMV AD169 which was amplified by PCR primers T85074-5. The  
 CC amplified product was digested with BglII and EcoRI and cloned into  
 CC BglII/EcoRI digested pB877 vector DNA to yield the plasmid pB877 UL105.  
 CC The amplified HCMV UL105 orf sequence and its translational product were  
 CC determined. The HCMV UL105 orf was recombined into the baculovirus AcNPV  
 CC genome by standard co-transfection protocols. The invention relates to a  
 CC new human cytomegalovirus (HCMV) primase. A screening assay for compounds  
 CC which inhibit HCMV primase (preferably with an IC50 of not greater than  
 CC 200nM), comprises incubation of the compound with the primase. The  
 CC primase of the invention is useful as a screening tool for HCMV  
 CC antivirals, and as a diagnostic tool for diseases resulting from HCMV  
 CC infection. 956 AA;  
 SQ Sequence 956 AA;  
 Query Match 80.0%; Score 44; DB 1; Length 956;  
 Best Local Similarity 62.5%; Pred. No. 2.43e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 359 TDLDFGDL 366  
 QY 139 SLSYGD 146  
 RESULT 6  
 ID W39372 standard; Protein; 1242 AA.  
 AC W39372;  
 DT 30-JUN-1999 (first entry)  
 DE Human IRS-1 and IRS-2 binding inhibitor protein.  
 KW Insulin receptor substrate; IRS-1; IRS-2; inhibitor; disease; obesity;  
 KW insulin resistance; 14-3-3 protein; inhibitor; screening; diabetes;  
 KW diabetic retinopathy; diabetic neuropathy; glucose tolerance;  
 KW hyperlipemia; hyperinsulinaemia; hyperlipemia; arteriosclerosis;  
 KW diabetic nephropathy; hyperinsulinaemia; hyperlipemia; arteriosclerosis;  
 KW peripheral embolism.  
 OS Homo sapiens.  
 PN WO9916462-A1.  
 PD 08-APR-1999.  
 PF 25-SEP-1998; JP-263719.  
 PR 29-SEP-1997; JP-263719.  
 PA (DAUC) DAICHI PHARM CO LTD.  
 PI Asano T, Kanda A, Kubo H, Yazaki Y;  
 DR WPI; 99-254929/21.  
 DR Treatment of insulin resistance using insulin receptor substrate  
 PT Binding inhibitor 16-22; 30pp; Japanese.  
 PS Disclosure; Page 16-22; 30pp; Japanese.  
 CC This invention describes a method for the treatment of diseases involving  
 CC insulin resistance using a substance which inhibits the binding of  
 CC insulin receptor substrate to 14-3-3 protein, identified by screening  
 CC potential inhibitors for their ability to inhibit this binding. The  
 CC composition described in the invention for the treatment of diseases  
 CC involving insulin resistance contains as active component, an inhibitor  
 CC of the binding of all or part of insulin receptor substrate 1 or 2  
 CC (IRS-1, IRS-2) to all or part of 14-3-3 protein. Disorders involving  
 CC insulin resistance include diabetes, diabetic retinopathy, diabetic  
 CC neuropathy, impaired glucose tolerance, diabetic nephropathy,  
 CC hyperinsulinaemia, hyperlipemia, arteriosclerosis, hypertension,  
 CC obesity, ischaemic heart disease, ischaemic brain disease and peripheral  
 CC embolism.  
 SQ Sequence. 1242 AA;  
 Query Match 80.0%; Score 44; DB 1; Length 1242;  
 Best Local Similarity 85.7%; Pred. No. 2.43e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Db 148 DLSYGDV 154  
 QY 140 DLSYGD 146

RESULT 7  
 ID R67708 standard; Protein; 1243 AA.  
 AC R67708;  
 DT 20-JUL-1995 (first entry)  
 DE Insulin receptor substrate-1.  
 KW Insulin receptor substrate-1; IRS-1;  
 OS non-insulin-dependent diabetes; NIDDM.  
 OS Homo sapiens.  
 PN WO9429345-A.  
 PD 22-DEC-1994.  
 PF 10-JUN-1994; DK0227.  
 PR 10-JUN-1993; DK-000683.  
 PR 09-AUG-1993; DK-000915.  
 PA (NOVO) NOVO-NORDISK AS.  
 PI Bjoerbaek C, Frederiksen KA, Pedersen O;  
 DR N-PSDB; Q75977.  
 DT DNA encoding mutated insulin receptor substrate 1 - and related  
 PT vectors, transformed cells and mutant proteins, useful for  
 PT detecting predisposition to non-insulin-dependent diabetes  
 PS Disclosure: Page 31-38; 58pp; English.  
 CC Genomic DNA from human leukocyte nuclei was subjected to PCR  
 CC amplification using primers (given in Q75978-Q76015) specific for the  
 CC human IRS-1 gene (Q75977). Amplified DNA was subjected to SSCP  
 CC analysis to detect mutations in the IRS-1 gene, indicating a  
 CC predisposition to NIDDM.  
 SQ Sequence 1243 AA;  
 Query Match 80.0%; Score 44; DB 1; Length 1243;  
 Best Local Similarity 85.7%; Pred. No. 2.43e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 DB 149 DLSYGDV 155  
 QY 140 DLSYGDL 146  
 RESULT 8  
 ID R28979 standard; protein; 99 AA.  
 AC R28979;  
 DT 24-MAR-1993 (first entry)  
 DE Bovine RFXBP.  
 KW Rapamycin; FK506; binding protein; RFXBP; prolyl isomerase;  
 KW immunosuppressant; cyclosporin A; macrolide; bovine; thymus; bRFXBP;  
 KW cis-trans prolyl isomerase activity; FKBP12.  
 OS Bos taurus.  
 PN WO9219745-A.  
 PD 12-NOV-1992.  
 PF 07-MAY-1992; U03993.  
 PR 08-MAY-1991; US-697113.  
 PA (VERT-) VERTEX PHARM INC.  
 PI Harding MW;  
 DR WPI; 92-398871/48.  
 DT New prolyl isomerase and rapamycin FK506 binding protein - useful  
 PT for screening potential immunosuppressive cpds.  
 PS Disclosure: Fig 1; 30pp; English.  
 CC This sequence corresponds to a fragment of a rapamycin FK506 binding  
 CC protein (RFXBP). RFXBP is a prolyl isomerase structurally related to  
 CC FK506 which does not bind the immunosuppressive cyclosporin A. RFXBP  
 CC binds FK506 and rapamycin with quantitatively significant selectivity.  
 CC RFXBP may be used in screening assays to detect new immunosuppressants  
 CC and to differentiate rapamycin-like cpds. from FK506-like cpds.  
 CC Rapamycin is a macrolide which is structurally related to FK506.  
 CC This RFXBP has been isolated from bovine thymus (bRFXBP) and was found  
 CC to be of low molecular weight, approx. 16,000, and to have cis-trans  
 CC prolyl isomerase activity. The N terminal of bRFXBP has been shown to  
 CC have over 50% homology to the N terminal of FKBP12.  
 SQ Sequence 99 AA;  
 Query Match 76.4%; Score 42; DB 1; Length 99;  
 Best Local Similarity 57.1%; Pred. No. 4.03e+02;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 DB 107 SELGYGE 113  
 QY 139 SDLSYGD 145  
 RESULT 9  
 ID R93551 standard; Protein; 141 AA.  
 AC R93551;  
 DT 25-JUN-1996 (first entry)  
 DE Human FKBP-13 immunophilin protein.  
 KW FKBP-13; immunophilin; FK506; rapamycin; rheumatoid arthritis;  
 KW diabetes; organ transplant; graft versus host disease;  
 OS immunosuppressant.  
 OS Homo sapiens.  
 FH Key  
 FT peptide  
 FT 1..21  
 FT /label= sig\_peptide  
 FT misc\_difference 22..61  
 FT /note= "corresponds to bovine thymus FKBP-13 N-  
 FT terminal sequence"  
 FT misc\_difference 117..120  
 FT /note= "endoplasmic reticulum retention sequence"  
 PN US5498597-A.  
 PD 12-MAR-1996.  
 PF 17-JAN-1992; US-822966.  
 PR 17-JAN-1992; US-822966.  
 PA (DAND) DANA FARBER CANCER INST INC.  
 PA (HARD) HARVARD COLLEGE.  
 PI Bierer BE, Burakoff SJ, Schreiber SL;  
 DR WPI; 96-159713/16.  
 DR N-PSDB; T18037.  
 DT Purified mammalian FKBP-13 polypeptide capable of binding FK506  
 PT useful for identifying and studying immunosuppressant drugs  
 PS Claim 1; Fig. 1; 12pp; English.  
 CC The FKBP-13 protein may be used for identifying immunosuppressant  
 CC drugs, and may be used in combination with immunosuppressant drugs  
 CC for therapeutic purposes in the treatment of autoimmune diseases e.g.  
 CC rheumatoid arthritis and type-I diabetes, organ transplant and  
 CC graft versus host disease. The recombinant form of the protein  
 CC could be potentially smaller and therefore easier to introduce  
 CC into cells than intact FKBP-13.  
 SQ Sequence 141 AA;  
 Query Match 76.4%; Score 42; DB 1; Length 141;  
 Best Local Similarity 57.1%; Pred. No. 4.03e+02;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 DB 107 SELGYGE 113  
 QY 139 SDLSYGD 145  
 RESULT 10  
 ID R28980 standard; Protein; 143 AA.  
 AC R28980;  
 DT 24-MAR-1993 (first entry)  
 DE bRFXBP.  
 KW Rapamycin; FK506; binding protein; RFXBP; prolyl isomerase;  
 KW immunosuppressant; cyclosporin A; macrolide; human; placental; bRFXBP;  
 KW cis-trans prolyl isomerase activity.  
 OS Homo sapiens.  
 PN WO9219745-A.  
 PD 12-NOV-1992.  
 PF 07-MAY-1992; U03993.  
 PR 08-MAY-1991; US-697113.  
 PA (VERT-) VERTEX PHARM INC.  
 PI Harding MW;  
 DR WPI; 92-398871/48.  
 DR N-PSDB; Q31004.  
 DT New prolyl isomerase and rapamycin FK506 binding protein - useful  
 PT for screening potential immunosuppressive cpds.  
 PS Disclosure: Fig 3; 30pp; English.  
 CC This sequence is a rapamycin FK506 binding protein (RFXBP). RFXBP is

CC a prolyl isomerase structurally related to FK506 which does not bind  
 CC the immunosuppressive cyclosporin A. FKBP binds FK506 and rapamycin  
 CC with quantitatively significant selectivity. FKBP may be used in  
 CC screening assays to detect new immunosuppressants and to differentiate  
 CC rapamycin-like cpds. from FK506-like cpds. Rapamycin is a macrolide  
 CC which is structurally related to FK506. This FKBP (hrFKBP) was found  
 CC to have a low molecular weight, approx. 13,200, and cis-trans prolyl  
 CC isomerase activity.  
 SQ Sequence 143 AA;

Query Match 76.4%; Score 42; DB 1; Length 143;  
 Best Local Similarity 57.1%; Pred. No. 4.03e-02;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Db 109 SELGYCE 115  
 QY 139 SDSLIGD 145

RESULT 11  
 ID W24593 standard; Protein; 186 AA.  
 AC W24593;  
 DT 04-AUG-1997 (first entry)  
 DE H. pylori cytoplasmic protein, 14572133.aa.  
 KW Transmembrane; cytoplasmic; cell envelope; flagella; transport;  
 KW secreted; periplasmic; chronic gastritis; duodenal ulcer disease;  
 KW activator; inhibitor; bacterial life cycle; vaccine; immunise;  
 KW detection; antisense; inhibition.  
 OS Helicobacter pylori.  
 FH Key Location/Qualifiers  
 FT misc\_difference 94 /note= "encoded by GGS"  
 FT misc\_difference 124 /note= "encoded by RCT"  
 FT misc\_difference 152 /note= "encoded by AAM"  
 FT misc\_difference 164 /note= "encoded by YCC"  
 FT misc\_difference 174 /note= "encoded by GRC"  
 FT misc\_difference 178 /note= "encoded by RCT"  
 FT misc\_difference 179 /note= "encoded by GRG"  
 FT misc\_difference 181 /note= "encoded by GAR"  
 FT misc\_difference 182 /note= "encoded by RGG"  
 FT W09719098-A1.  
 PD 29-MAY-1997.  
 PF 15-NOV-1996; U18542.  
 PR 17-NOV-1995; US-561469.  
 PA (ASTR ) ASTRA AB.  
 PI Smith DH;  
 DR WPI: 97-298052/27.  
 DR N-PSDB: T77411.  
 PT Helicobacter pylori nucleic acid sequences and related proteins -  
 PT used for diagnostics and therapeutics  
 PS Claim 18; Page 136; 235pp; English.  
 CC This sequence represents an H. pylori cytoplasmic protein.  
 CC Helicobacter pylori has been strongly linked to chronic gastritis and  
 CC duodenal ulcer disease. The nucleic acid sequences of the invention  
 CC are used to evaluate compounds, especially activators or inhibitors of  
 CC bacterial life cycle, for the ability to bind an H. pylori nucleic acid  
 CC sequence. The nucleic acid sequences, and corresponding proteins, are  
 CC also useful for generating vaccines for immunising subjects against H.  
 CC pylori or for use in detecting the presence of Helicobacter species in  
 CC a sample. Antisense nucleic acid sequences of these sequences are  
 CC used to inhibit expression of a gene from Helicobacter species. H.  
 CC pylori whole genomic DNA was isolated and nebulised to a median size of  
 CC 2000 bp. Purified DNA fragments were blunt-ended and ligated to unique  
 CC BstXI-linker adapters in 100-1000 fold molar excess. These linkers are  
 CC complementary to the BstXI-cut PMPX vectors, while the overhang is not

CC self-complementary. Therefore the linkers will not concatamerise nor  
 CC will the cut vector re-ligate itself easily. The linker-adaptor inserts  
 CC were ligated to each of the 20 PMPX vectors to construct a series of  
 CC shotgun subclone libraries. The purified DNA samples were then  
 CC sequenced.  
 CC Note: the ORF/protein reference number for this sequence was obtained  
 CC from the related specification, WO9640893.  
 SQ Sequence 186 AA;

Query Match 76.4%; Score 42; DB 1; Length 186;  
 Best Local Similarity 62.5%; Pred. No. 4.03e-02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 126 TELSUGVL 133  
 QY 139 SDSLIGDL 146

RESULT 12  
 ID W20146 standard; Protein; 187 AA.  
 AC W20146; 1997 (first entry)  
 DT 29-JUL-1997  
 DE H. pylori cytoplasmic protein, 14572133.aa.  
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacterium; life cycle; activator;  
 KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;  
 KW diagnosis.  
 OS Helicobacter pylori.  
 FH Key Location/Qualifiers  
 FT misc\_difference 94 /note= "encoded by GGS"  
 FT misc\_difference 124 /label= Unknown  
 FT /note= "encoded by RCT"  
 FT misc\_difference 132 /label= Unknown  
 FT /note= "encoded by AAM"  
 FT misc\_difference 164 /label= Unknown  
 FT /note= "encoded by YCC"  
 FT misc\_difference 174 /label= Unknown  
 FT /note= "encoded by GRC"  
 FT misc\_difference 178 /label= Unknown  
 FT /note= "encoded by RCT"  
 FT misc\_difference 179 /label= Unknown  
 FT /note= "encoded by GRG"  
 FT misc\_difference 181 /note= "encoded by GAR"  
 FT misc\_difference 182 /label= Unknown  
 FT /note= "encoded by RGG"  
 FT W09640893-A1.  
 PD 19-DEC-1996.  
 PF 06-JUN-1996; U09122.  
 PR 07-JUN-1995; US-487032.  
 PR 01-APR-1996; US-630405.  
 PA (ASTR ) ASTRA AB.  
 PI Berglundh OT, Smith D, Mellgaard BL;  
 DR WPI: 97-052306/05.  
 DR N-PSDB: T67731.  
 PT Helicobacter pylori nucleic acid sequences and related  
 PT polypeptides) - useful for vaccines to treat or prevent H. pylori  
 PT infection, and to detect Helicobacter  
 PS Claim 61; Page 368; 1481pp; English.  
 CC This sequence represents a H. pylori cytoplasmic protein.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors.  
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
 CC overlapping contigs generated by mechanically shearing the bacterial

CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,  
 CC and the predicted coding regions defined by computer evaluation. To  
 CC identify likely H. pylori antigens for vaccine development, the amino  
 CC acid sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.  
 SQ Sequence 187 AA;

Query Match 76.4%; Score 42; DB 1; Length 187;  
 Best Local Similarity 62.5%; Pred. No. 4.03e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 126 TELSXYGL 133  
 :|||||  
 QY 139 SDLSYGL 146

## RESULT 13

ID Y10982 standard; Protein; 248 AA.

AC Y10982;  
 DT 08-JUN-1999 (first entry)  
 DE H. pylori ORF 05ep10815.16131925\_c2.97 outer membrane protein.  
 KW Vaccine: probe; diagnostic; ORF: cell envelope protein;  
 KW secreted protein; cellular protein.  
 OS Helicobacter pylori.  
 PN W09818323-AL.

PD 07-MAY-1998.

PF 28-OCT-1997; U19575.

PR 14-JUL-1997; US-891928.

PR 28-OCT-1996; US-739150.

PR 06-DEC-1996; US-759739.

PA (ASTR ) ASTRA AB.

PI Alm RA, Smith D.

DR WPI; 98-271811/24.

DR N-PSDB; X30449.

PT Helicobacter pylori nucleic acids and proteins - used to develop  
 PT products for the detection, prevention and treatment of H. pylori  
 PT infections

PS Claims 27, 31; Page 186; 279pp; English.

CC Recombinant or substantially pure preparations of H. pylori polypeptides  
 CC are disclosed, together with the nucleic acids encoding them. In all,  
 CC 73 ORFs are shown. The proteins are variously cell envelope proteins,  
 CC secreted proteins or other cellular proteins. Vaccines containing the  
 CC nucleic acids or proteins are claimed, as are probes containing at least  
 CC 8 nucleotides from the nucleic acid sequences. The vaccines are useful  
 CC for treating or reducing the risk of H. pylori infections, and the  
 CC probes can be used diagnostically for detecting the presence of  
 CC Helicobacter in a sample. The products are also of use in screening  
 CC for compounds having the ability to interfere with the H. pylori life  
 CC cycle or to inhibit H. pylori infection.  
 SQ Sequence 248 AA;

Query Match 76.4%; Score 42; DB 1; Length 248;  
 Best Local Similarity 62.5%; Pred. No. 4.03e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 128 TELSXYGL 135  
 :|||||  
 QY 139 SDLSYGL 146

## RESULT 14

ID W20740 standard; Protein; 253 AA.

AC W20740;

DE 16-JUN-1997 (first entry)

DE H. pylori cytoplasmic protein, 06ee10207orf2.

KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;

KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;

KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.

OS Helicobacter pylori.

PN W09640893-AL.

PD 19-DEC-1996.

PF 06-JUN-1996; U09122.

PR 07-JUN-1995; US-487032.

PR 01-APR-1996; US-630405.

PA (ASTR ) ASTRA AB.

PI Berglindh OT, Smith D, Mellgaard BL;

DR WPI; 97-052306/05.

DR N-PSDB; T67993.

PT Helicobacter pylori nucleic acid sequences and related  
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
 PT infection, and to detect Helicobacter.

PS Claim 61; Page 1156; 1481pp; English.

CC The present sequence is a Helicobacter pylori cytoplasmic protein.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors.

CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
 CC overlapping contigs generated by mechanically shearing the bacterial  
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,  
 CC and the predicted coding regions defined by computer evaluation. To  
 CC identify likely H. pylori antigens for vaccine development, the amino  
 CC acid sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.  
 SQ Sequence 253 AA;

Query Match 76.4%; Score 42; DB 1; Length 253;

Best Local Similarity 62.5%; Pred. No. 4.03e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 176 TELSXYGL 183  
 :|||||  
 QY 139 SDLSYGL 146

## RESULT 15

ID W21783 standard; Protein; 275 AA.

AC W21783;

DT 11-MAR-1998 (first entry)

DE Protein encoded by ORF F from the GS region of M. tuberculosis.

KW GS; pathogenicity island; pathogenic protein; mycobacterial disease;

KW cellular recognition receptor; pathogenic mycobacteria; Crohn's disease;

KW vaccine; inflammatory disease; sarcoidosis; Johne's disease; ss.

OS Mycobacterium tuberculosis.

PN W09723624-42.

PD 03-JUL-1997.

PF 23-DEC-1996; G03221.

PR 21-DEC-1995; GB-026178.

PA (SGEO-) ST GEORGE'S HOSPITAL MEDICAL SCHOOL.

PI Doran T, Ford J, Hermon-Taylor J, Loughlin M, Millar D;

PI Sumar N, Tizard M;

DR WPI; 97-351061/32.

DR N-PSDB; T74477.

PT New isolated pathogenicity island from mycobacteria - used to  
 PT develop products for detection, diagnosis, prevention and treatment  
 PT of mycobacteria infections

PS Claim 1; Page:56; 62pp; English.

CC The present sequence represents the protein encoded by open reading frame  
 CC (ORF) F, from M. tuberculosis. This ORF F has been found to have homology  
 CC with ORF F of a novel polynucleotide sequence designated "GS". GS is a  
 CC pathogenicity island of 8 kb of DNA comprising a core region of 5.75 kb  
 CC with multiple ORFs and an adjacent transmissible element of 2.5 kb. The  
 CC ORFs, and also the transmissible element, encode proteins which may be  
 CC linked to pathogenicity, such as providing receptors for cellular  
 CC recognition. GS was discovered and characterised using differential DNA  
 CC analysis technology. It is found within Mycobacterium paratuberculosis  
 CC and it has also been identified in Mycobacterium avium subspecies  
 CC silvaticum. These pathogenic mycobacteria cause chronic inflammation of  
 CC the intestine and Crohn's disease in humans. The protein products of the  
 CC ORFs of GS can be used for detecting mycobacteria or for diagnosing,  
 CC treating or preventing mycobacterial disease. In particular they can be

US-09-376-430-2-16.rag

Thu May 11 06:50:00 2000

CC used as vaccines for inflammatory diseases such as Crohn's disease or  
CC sarcoidosis in humans or Johne's disease in animals.  
SQ Sequence 275 AA;

Query Match 76.4%; Score 42; DB 1; Length 275;  
Best Local Similarity 75.0%; Pred. NO. 4.03e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Db 135 SLDVYGDV 142  
QY 139 SLDVYGDV 146

Search completed: Wed May 10 13:15:34 2000  
Job time : 8 secs.





```
CC TELEFAX: 516-742-4366
CC TELEX: 230 901 SANS UR
CC INFORMATION FOR SEQ ID NO: 12:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 507 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 507 AA; 53491 MW; 1252534 CN;

Query Match      87.3%; Score 48; DB 2; Length 507;
Best Local Similarity 85.7%; Pred. No. 4.96e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 30 DLGYGDL 36
   ||:||||
QY 140 DLSYGDL 146

RESULT 2
ID US-08-484-494-12 STANDARD; PRT; 507 AA.
XX
AC xxxxxx
XX
DT
DT
XX
XX
DE
XX
Sequence 12, Application US/08484494
Sequence 12, Application US/08484494
Patent No. 5798239
GENERAL INFORMATION:
APPLICANT: Wilson, Peter J
APPLICANT: Morris, Charles P
APPLICANT: Anson, Donald S
APPLICANT: Occhiodoro, Teresa
APPLICANT: Bielicki, Julie
APPLICANT: Clements, Peter R
APPLICANT: Hopwood, John J
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
TITLE OF INVENTION: IDURONATE 2-SULFATASE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,494
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 84162
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 507 AA; 53491 MW; 1252534 CN;
SQ
```

```
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 507 AA; 53491 MW; 1252534 CN;

Query Match      87.3%; Score 48; DB 1; Length 507;
Best Local Similarity 85.7%; Pred. No. 4.96e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 30 DLGYGDL 36
   ||:||||
QY 140 DLSYGDL 146

RESULT 3
ID US-08-484-493-12 STANDARD; PRT; 507 AA.
XX
AC xxxxxx
XX
DT
DT
XX
XX
DE
XX
Sequence 12, Application US/08484493
Sequence 12, Application US/08484493
Patent No. 5728381
GENERAL INFORMATION:
APPLICANT: Wilson, Peter J
APPLICANT: Morris, Charles P
APPLICANT: Anson, Donald S
APPLICANT: Occhiodoro, Teresa
APPLICANT: Bielicki, Julie
APPLICANT: Clements, Peter R
APPLICANT: Hopwood, John J
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
TITLE OF INVENTION: IDURONATE 2-SULFATASE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,493
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 84162
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 507 AA; 53491 MW; 1252534 CN;
SQ
```

Query Match 87.3%; Score 48; DB 1; Length 507;







US-09-376-430-2-16.rai

Thu May 11 06:50:00 2000

CC	APPLICANT: Garcia, Bruno D.	OTHER INFORMATION: Unidentified at time of filing	CC
CC	APPLICANT: Gutierrez, Santiago	FEATURE: NAME/KEY: unsure	CC
CC	APPLICANT: Barredo, Jose L.	LOCATION: (181)	CC
CC	APPLICANT: Von Doehren, Hans	OTHER INFORMATION: Unidentified at time of filing	CC
CC	APPLICANT: Palissa, Harriet	FEATURE: NAME/KEY: unsure	CC
CC	APPLICANT: Van Liempt, Henk	LOCATION: (438)	CC
CC	APPLICANT: Montenegro, Eduardo P.	OTHER INFORMATION: Unidentified at time of filing	CC
CC	TITLE OF INVENTION: A Method for Influencing Beta-Lactam	FEATURE: NAME/KEY: unsure	CC
CC	TITLE OF INVENTION: Antibiotic Production and for Isolation of Large	LOCATION: (439)	CC
CC	TITLE OF INVENTION: Quantities of ACV Synthetase	OTHER INFORMATION: Unidentified at time of filing	CC
CC	NUMBER OF SEQUENCES: 27	FEATURE: NAME/KEY: unsure	CC
CC	CORRESPONDENCE ADDRESS:	LOCATION: (567)	CC
CC	ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff	OTHER INFORMATION: Unidentified at time of filing	CC
CC	STREET: 300 South Wacker Drive	FEATURE: NAME/KEY: unsure	CC
CC	CITY: Chicago	LOCATION: (568)	CC
CC	STATE: Illinois	OTHER INFORMATION: Unidentified at time of filing	CC
CC	COUNTRY: USA	FEATURE: NAME/KEY: unsure	CC
CC	ZIP: 60606	LOCATION: (571)	CC
CC	COMPUTER READABLE FORM:	OTHER INFORMATION: Unidentified at time of filing	CC
CC	MEDIUM TYPE: Floppy disk	FEATURE: NAME/KEY: unsure	CC
CC	COMPUTER: IBM PC compatible	LOCATION: (575)	CC
CC	OPERATING SYSTEM: PC-DOS/MS-DOS	OTHER INFORMATION: Unidentified at time of filing	CC
CC	SOFTWARE: Patent in Release #1.0, Version #1.25	FEATURE: NAME/KEY: unsure	CC
CC	CURRENT APPLICATION DATA:	LOCATION: (580)	CC
CC	APPLICATION NUMBER: US/08/222,617A	OTHER INFORMATION: Unidentified at time of filing	CC
CC	FILING DATE: 04-APR-1994	FEATURE: NAME/KEY: unsure	CC
CC	CLASSIFICATION: 435	LOCATION: (589)	CC
CC	ATTORNEY/AGENT INFORMATION:	OTHER INFORMATION: Unidentified at time of filing	CC
CC	REFERENCE/DOCKET NUMBER: 97,157	FEATURE: NAME/KEY: unsure	CC
CC	INFORMATION FOR SEQ ID NO: 2:	LOCATION: (593)	CC
CC	SEQUENCE CHARACTERISTICS:	OTHER INFORMATION: Unidentified at time of filing	CC
CC	LENGTH: 3778 amino acids	FEATURE: NAME/KEY: unsure	CC
CC	TYPE: amino acid	LOCATION: (599)	CC
CC	TOPOLOGY: linear	OTHER INFORMATION: Unidentified at time of filing	CC
CC	MOLECULE TYPE: protein	FEATURE: NAME/KEY: unsure	CC
CC	SEQUENCE 3778 AA; 424374 MW; 73326696 CN;	LOCATION: (600)	CC
CC		OTHER INFORMATION: Unidentified at time of filing	CC
CC		FEATURE: NAME/KEY: unsure	CC
CC		LOCATION: (608)	CC
CC		OTHER INFORMATION: Unidentified at time of filing	CC
CC		FEATURE: NAME/KEY: unsure	CC
CC		LOCATION: (610)	CC
CC		OTHER INFORMATION: Unidentified at time of filing	CC
CC		FEATURE: NAME/KEY: unsure	CC
CC		LOCATION: (611)	CC
CC		OTHER INFORMATION: Unidentified at time of filing	CC
CC		FEATURE: NAME/KEY: unsure	CC
CC		LOCATION: (621)	CC
CC		OTHER INFORMATION: Unidentified at time of filing	CC
CC		FEATURE: NAME/KEY: unsure	CC
CC		LOCATION: (624)	CC
CC		OTHER INFORMATION: Unidentified at time of filing	CC
CC		FEATURE: NAME/KEY: unsure	CC
CC		LOCATION: (625)	CC
CC		OTHER INFORMATION: Unidentified at time of filing	CC

Db	347 ELTYGEL 353	
QY	140 DLSYGD 146	
	Query Match 85.5%; Score 47; DB 2; Length 3778;	
	Best Local Similarity 57.1%; Pred. No. 6.43e+01;	
	Mismatches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;	
RESULT		
ID	US-08-956-242-4	STANDARD: PRT: 888 AA.
XX	xxxxxx	
XX		
DT		
XX		
DE	Sequence 4, Application US/08956242C	
XX		
CC	Sequence 4, Application US/08956242C	
CC	Patent No. 5986081	
CC	GENERAL INFORMATION:	
CC	APPLICANT: Ganetsky, Barry S.	
CC	APPLICANT: Titus, Steven A.	
CC	TITLE OF INVENTION: Polynucleotides Encoding Herg-3	
CC	FILE REFERENCE: 960296.94550	
CC	CURRENT APPLICATION NUMBER: US/08/956,242C	
CC	CURRENT FILING DATE: 1997-10-22	
CC	NUMBER OF SEQ ID NOS: 13	
CC	SOFTWARE: Patent in Ver. 2.0	
CC	SEQ ID NO 4	
CC	LENGTH: 888	
CC	TYPE: PRT	
CC	ORGANISM: Homo sapien	
CC	FEATURE: NAME/KEY: unsure	
CC	LOCATION: (133)	





```

CC Patent No. 5763590
CC GENERAL INFORMATION:
CC APPLICANT: Peattie, Debra A.
CC APPLICANT: Harding, Matthew W.
CC APPLICANT: Livingston, David J.
CC TITLE OF INVENTION: ISOLATION OF AN Mr 52,000 FK506 BINDING
CC TITLE OF INVENTION: PROTEIN AND MOLECULAR CLONING OF A CORRESPONDING HUMAN
CC TITLE OF INVENTION: CDNA
CC NUMBER OF SEQUENCES: 32
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith and Reynolds, P.C.
CC STREET: Two Militia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/336,618
CC FILING DATE: 09-NOV-1994
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/963,325
CC FILING DATE: 16-OCT-1992
CC APPLICATION NUMBER: US 07/777,752
CC FILING DATE: 11-OCT-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/
CC FILING DATE: 09-OCT-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: VP191-06A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 142 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 142 AA; 15649 MW; 106171 CN;
SQ
Query Match 76.4%; Score 42; DB 1; Length 142;
Best Local Similarity 57.1%; Pred. No. 2.29e-02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 108 SELGYE 114
Qy 139 SDSLXGD 145

Search completed: Wed May 10 13:16:44 2000
Job time : 53 secs.

```

\*\*\*\*\*  
 W P S R L H  
 \*\*\*\*\*  
 (TR)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed May 10 12:34:42 2000; MasPar time 44.05 Seconds  
 3.530 Million cell updates/sec  
 Tabular output not generated.

Title: >US-09-376-430-2  
 Description: (62-73) from US09376430A.pep (11 of 25)  
 Perfect Score: 87  
 Sequence: 1 FNGDEAYDOCTN 12

Scoring table: PAM 150  
 Gap 11

Searched: 131253 seqs, 12956647 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: a-issued  
 1:5A\_COMB 2:5B\_COMB 3:PCT\_COMB 4:backfiles1

Statistics: Mean 16.446; Variance 49.600; scale 0.332

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	53	60.9	170	1 US-08-650-	Sequence 101, Applicat	2.88e+01
2	50	57.5	264	2 US-07-857-	Sequence 34, Applicatio	6.10e+01
3	48	55.2	178	1 US-08-650-	Sequence 2, Applicatio	1.00e+02
4	48	55.2	241	4 5213972-7	Patent No. 5213972	1.00e+02
5	47	54.0	411	2 US-08-338-	Sequence 3, Applicatio	1.28e+02
6	47	54.0	1230	2 US-08-968-	Sequence 35, Applicatio	1.28e+02
7	45	51.7	17	1 US-08-137-	Sequence 144, Applicat	2.07e+02
8	45	51.7	17	1 US-08-436-	Sequence 144, Applicat	2.07e+02
9	45	51.7	116	3 PCT-US95-0	Sequence 41, Applicati	2.07e+02
10	45	51.7	116	2 US-08-561-	Sequence 41, Applicati	2.07e+02
11	45	51.7	135	2 US-08-436-	Sequence 100, Applicat	2.07e+02
12	45	51.7	135	1 US-08-137-	Sequence 100, Applicat	2.07e+02
13	45	51.7	135	1 US-08-137-	Sequence 102, Applicat	2.07e+02
14	45	51.7	135	1 US-08-137-	Sequence 112, Applicat	2.07e+02
15	45	51.7	135	2 US-08-137-	Sequence 27, Applicati	2.07e+02
16	45	51.7	135	2 US-08-436-	Sequence 102, Applicat	2.07e+02
17	45	51.7	135	2 US-08-436-	Sequence 112, Applicat	2.07e+02
18	45	51.7	237	3 PCT-US95-1	Sequence 5, Applicatio	2.07e+02
19	45	51.7	237	2 US-08-468-	Sequence 5, Applicatio	2.07e+02
20	45	51.7	307	1 US-08-713-	Sequence 1, Applicatio	2.07e+02
21	45	51.7	307	2 US-08-919-	Sequence 1, Applicatio	2.07e+02
22	45	51.7	335	2 US-08-875-	Sequence 3, Applicatio	2.07e+02
23	45	51.7	340	1 US-08-606-	Sequence 5, Applicatio	2.07e+02

24	45	51.7	410	1 US-08-579-	Sequence 8, Applicatio	2.07e+02
25	45	51.7	410	1 US-08-579-	Sequence 2, Applicatio	2.07e+02
26	45	51.7	410	1 US-08-579-	Sequence 6, Applicatio	2.07e+02
27	45	51.7	649	1 US-07-894-	Sequence 2, Applicatio	2.07e+02
28	45	51.7	650	1 US-07-893-	Sequence 1, Applicatio	2.07e+02
29	45	51.7	705	2 US-08-023-	Sequence 19, Applicati	2.07e+02
30	45	51.7	705	2 US-08-663-	Sequence 19, Applicati	2.07e+02
31	45	51.7	705	3 PCT-US95-1	Sequence 19, Applicati	2.07e+02
32	45	51.7	705	2 US-08-362-	Sequence 19, Applicati	2.07e+02
33	45	51.7	748	3 PCT-US91-0	Sequence 28, Applicati	2.07e+02
34	45	51.7	748	3 PCT-US91-0	Sequence 6, Applicatio	2.07e+02
35	45	51.7	748	3 PCT-US93-0	Sequence 4, Applicatio	2.07e+02
36	45	51.7	748	3 PCT-US91-0	Sequence 34, Applicati	2.07e+02
37	45	51.7	748	3 PCT-US91-0	Sequence 24, Applicati	2.07e+02
38	45	51.7	1454	3 PCT-US93-0	Sequence 16, Applicati	2.07e+02
39	45	51.7	1454	3 PCT-US93-0	Sequence 45, Applicati	2.07e+02
40	45	51.7	1454	3 PCT-US93-0	Sequence 8, Applicatio	2.07e+02
41	45	51.7	1454	3 PCT-US91-0	Sequence 42, Applicati	2.07e+02
42	45	51.7	1454	3 PCT-US93-0	Sequence 2, Applicatio	2.07e+02
43	45	51.7	1454	3 PCT-US93-0	Sequence 47, Applicati	2.07e+02
44	45	51.7	1454	3 PCT-US93-0	Sequence 44, Applicati	2.07e+02
45	45	51.7	1454	3 PCT-US91-0	Sequence 26, Applicati	2.07e+02

## ALIGNMENTS

RESULT	1				
ID	US-08-650-528-101	STANDARD;	PRT;	170 AA.	
XX	xxxxxx				
DT					
XX					
DE	Sequence 101, Application US/08650528				
CC	Patent No. 5780278				
CC	GENERAL INFORMATION:				
CC	APPLICANT: MILLER, GERALDINE G.				
CC	APPLICANT: PEEK, JR., RICHARD M.				
CC	APPLICANT: THOMPSON, STUART A.				
CC	APPLICANT: BLASER, MARTIN J.				
CC	TITLE OF INVENTION: ICEA GENE AND RELATED METHODS				
CC	NUMBER OF SEQUENCES: 101				
CC	CORRESPONDENCE ADDRESS:				
CC	ADDRESS: NEEDLE & ROSENBERG, P.C.				
CC	STREET: Suite 1200, 127 Peachtree Street				
CC	CITY: Atlanta				
CC	STATE: Georgia				
CC	COUNTRY: USA				
CC	ZIP: 30303				
CC	COMPUTER READABLE FORM:				
CC	MEDIUM TYPE: Floppy disk				
CC	COMPUTER: IBM PC compatible				
CC	OPERATING SYSTEM: PC-DOS/MS-DOS				
CC	SOFTWARE: Patentin Release #1.0, Version #1.30				
CC	CURRENT APPLICATION DATA:				
CC	APPLICATION NUMBER: US/08/650.528				
CC	FILING DATE:				
CC	CLASSIFICATION: 435				
CC	ATTORNEY/AGENT INFORMATION:				
CC	NAME: Spratt, Gwendolyn D.				
CC	REGISTRATION NUMBER: 36,016				
CC	REFERENCE/DOCKET NUMBER: 22000.0049				
CC	TELECOMMUNICATION INFORMATION:				
CC	TELEPHONE: 404/688-0770				
CC	TELEFAX: 404/688-9880				
CC	INFORMATION FOR SEQ ID NO: 101:				
CC	SEQUENCE CHARACTERISTICS:				
CC	LENGTH: 170 amino acids				
CC	TYPE: amino acid				
CC	STRANDEDNESS: not relevant				
CC	TOPOLOGY: linear				

```

Db 197 FHGHDNYDQ 205
   I : I : I I I
Qy 62 FNGDEAYDQ 70

RESULT 3
ID US-08-650-528-2 STANDARD; PRT; 178 AA.
XX
XX
XX
XX
XX
XX
DE Sequence 2, Application US/08650528
XX
XX Sequence 2, Application US/08650528
CC Patent No. 5780278
CC GENERAL INFORMATION: 't
CC APPLICANT: MILLER, GERALDINE G.
CC APPLICANT: PEEK, JR., RICHARD M.
CC APPLICANT: THOMPSON, STUART A.
CC APPLICANT: BLASER, MARTIN J.
CC TITLE OF INVENTION: IDE A GENE AND RELATED METHODS
CC NUMBER OF SEQUENCES: 101
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: NEEDLE & ROSENBERG, P.C.
CC STREET: Suite 1200, 127 Peachtree Street
CC City: Atlanta
CC STATE: Georgia
CC COUNTRY: USA
CC ZIP: 30303
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/650,528
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Spratt, Gwendolyn D.
CC REGISTRATION NUMBER: 36,016
CC REFERENCE/DOCKET NUMBER: 22000.0049
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 404/688-0770
CC TELEFAX: 404/688-9880
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 178 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 178 AA; 20615 MW; 154543 CN;

Query Match 55.28; Score 48; DB 1; Length 178;
Best Local Similarity 40.0%; Pred. No. 1.00e+02;
Matches 4; Conservative 3; Mismatches 3; Indels

Db 127 YEGERAYDGC 136
   I : I : I I I
Qy 62 FNGDEAYDQ 71

RESULT 4
ID ID 5213972-7 STANDARD; PRT; 261 AA.
XX
XX
XX
XX
XX
XX
XX 01-JAN-1900
XX
XX Patent No. 5213972.

```



CC Patent No. 5213972  
CC APPLICANT: MCCANDLISS, RUSSELL J.; ANDERSON, DAVID M.  
CC TITLE OF INVENTION: FERMENTATION PROCESS FOR THE PRODUCTION  
CC OF PYRIMIDINE DEOXYRIBONUCLEOSIDES  
CC NUMBER OF SEQUENCES: 9  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/07/448,158  
CC FILING DATE: 08-DEC-1989  
CC SEQ ID NO: 7  
CC LENGTH: 241  
CC SEQUENCE 261 AA; 30594 MW; 391095 CN;  
SQ  
Query Match 55.28; Score 48; DB 4; Length 241;  
Best Local Similarity 40.0%; Pred. No. 1.00e+02;  
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
Db 94 YDNDEFYDC 103  
QY 62 FNGDEAYDQC 71  
RESULT 5  
ID US-08-338-530A-3 STANDARD; PRT; 411 AA.  
XX  
XX  
XX  
XX  
XX  
XX  
XX  
XX  
XX  
XX  
XX  
Sequence 3, Application US/08338530A  
CC  
XX  
XX  
Sequence 3, Application US/08338530A  
CC  
Patent No. 5922327  
CC  
GENERAL INFORMATION:  
CC APPLICANT: CRABB, Brendan S.  
CC APPLICANT: STUDDERT, Michael J.  
CC TITLE OF INVENTION: EQUINE HERPESVIRUS GLYCOPROTEINS  
CC NUMBER OF SEQUENCES: 13  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Foley & Lardner  
CC STREET: 3000 K Street, N.W., Suite 500  
CC CITY: Washington  
CC STATE: D.C.  
CC COUNTRY: USA  
CC ZIP: 20007-5109  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: IBM PC compatible  
CC SOFTWARE: Patent in Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/338,530A  
CC FILING DATE: 25-JAN-1995  
CC CLASSIFICATION: 424  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/AU93/00253  
CC FILING DATE: 28-MAY-1993  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: BENT, Stephen A.  
CC REGISTRATION NUMBER: 29,768  
CC REFERENCE/DOCKET NUMBER: 40268/120/CSMB  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (202)672-5300  
CC TELEFAX: (202)672-5399  
CC TELEX: 904136  
CC INFORMATION FOR SEQ ID NO: 3:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 411 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
SQ SEQUENCE 411 AA; 45269 MW; 858965 CN;  
Query Match 54.0%; Score 47; DB 2; Length 411;

Best Local Similarity 75.0%; Pred. No. 1.28e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Db 286 FVGDEYD 293  
QY 62 FNGDEAYD 69  
RESULT 6  
ID US-08-968-542C-35 STANDARD; PRT; 1230 AA.  
XX  
XX  
XX  
XX  
XX  
XX  
XX  
XX  
XX  
XX  
XX  
Sequence 35, Application US/08968542C  
CC  
XX  
XX  
Sequence 35, Application US/08968542C  
CC  
Patent No. 5981728  
CC  
GENERAL INFORMATION:  
CC APPLICANT: Myers, et al.  
CC TITLE OF INVENTION: dulla Codes For A No. 5981728el Starch  
CC TITLE OF INVENTION: Synthase  
CC NUMBER OF SEQUENCES: 35  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: McGregor & Adler, LLP  
CC STREET: 8011 Candle Lane  
CC CITY: Houston  
CC STATE: TX  
CC COUNTRY: USA  
CC ZIP: 77071  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: 3.5 floppy disk  
CC COMPUTER: Apple Macintosh  
CC OPERATING SYSTEM: Macintosh  
CC SOFTWARE: Microsoft Word 6.0.1 for Macintosh  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/968,542C  
CC FILING DATE: NO. 5981728ember 12, 1997  
CC CLASSIFICATION: 800  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER:  
CC FILING DATE:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Benjamin Aaron Adler, Ph.D., J.D.  
CC REGISTRATION NUMBER: 35,423  
CC REFERENCE/DOCKET NUMBER: D6036  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (713) 777-2321  
CC TELEFAX: (713) 777-6908  
CC INFORMATION FOR SEQ ID NO: 35:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1230 amino acid residues  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE:  
CC DESCRIPTION: amino acid  
CC HYPOTHETICAL: no  
CC ANTI-SENSE: no  
CC FRAGMENT TYPE:  
SQ SEQUENCE 1230 AA; 139110 MW; 7670544 CN;  
Query Match 54.0%; Score 47; DB 2; Length 1230;  
Best Local Similarity 62.5%; Pred. No. 1.28e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Db 411 FNGODVYD 418  
QY 62 FNGDEAYD 69  
RESULT 7  
ID US-08-137-117D-144 STANDARD; PRT; 17 AA.



US-09-376-430-2-11.ra1

Thu May 11 06:49:48 2000

```

XX Sequence 41, Application PC/TUS9501219
DE
XX Sequence 41, Application PC/TUS9501219
CC GENERAL INFORMATION:
CC APPLICANT: Bendig, Mary M.
CC APPLICANT: Leger, Olivier J.
CC APPLICANT: Saldanha, Jose
CC APPLICANT: Jones, S. Tarran
CC TITLE OF INVENTION: Humanized Antibodies Against Leukocyte
CC TITLE OF INVENTION: Adhesion Molecule VLA-4
CC NUMBER OF SEQUENCES: 45
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourie and Crew
CC STREET: One Market Plaza, Steuart Tower, Suite 2000
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94105
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/01219
CC FILING DATE: 25-JAN-1995
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/186,269
CC FILING DATE: 25-JAN-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Smith, William L.
CC REGISTRATION NUMBER: 30,223
CC REFERENCE/DOCKET NUMBER: 15270-14
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-543-9600
CC TELEFAX: 415-543-5043
CC INFORMATION FOR SEQ ID NO: 41:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 116 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 116 AA; 12703 MW; 79291 CN;

Query Match 51.7%; Score 45; DB 3; Length 116;
Best Local Similarity 55.6%; Pred. No. 2.07e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 54 FNGGTSYNO 62
QY 62 FNGDEAYDQ 70

RESULT 10
ID US-08-561-521-41 STANDARD; PRT; 116 AA.
XX
XX
AC xxxxxx
XX
XX
DT
XX
DE
XX
CC Sequence 41, Application US/08561521
CC Patent No. 5840299
CC GENERAL INFORMATION:
CC APPLICANT: Bendig, Mary M.
CC APPLICANT: Leger, Olivier J.
CC APPLICANT: Saldanha, Jose
CC APPLICANT: Jones, S. Tarran
CC TITLE OF INVENTION: Humanized Antibodies Against Leukocyte

```

```

CC TITLE OF INVENTION: Adhesion Molecule VLA-4
CC NUMBER OF SEQUENCES: 45
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourie and Crew
CC STREET: One Market Plaza, Steuart Tower, Suite 2000
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94105
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/561,521
CC FILING DATE:
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/08/186,269A
CC FILING DATE: 25-JAN-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Smith, William L.
CC REGISTRATION NUMBER: 30,223
CC REFERENCE/DOCKET NUMBER: 15270-14
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-543-9600
CC TELEFAX: 415-543-5043
CC INFORMATION FOR SEQ ID NO: 41:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 116 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 116 AA; 12703 MW; 79291 CN;

Query Match 51.7%; Score 45; DB 2; Length 116;
Best Local Similarity 55.6%; Pred. No. 2.07e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 54 FNGGTSYNO 62
QY 62 FNGDEAYDQ 70

RESULT 11
ID US-08-436-717-100 STANDARD; PRT; 135 AA.
XX
XX
AC xxxxxx
XX
XX
DT
XX
DE
XX
CC Sequence 100, Application US/08436717
CC Patent No. 5817790
CC GENERAL INFORMATION:
CC APPLICANT: TSUCHIYA, Masayuki
CC APPLICANT: SATO, Koh
CC APPLICANT: BENDIG, Mary
CC APPLICANT: JONES, Steven
CC APPLICANT: SALDANHA, Jose
CC TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
CC TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR
CC NUMBER OF SEQUENCES: 158
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 3000 K Street, N.W., Suite 500
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: USA
CC ZIP: 20007-5109

```

CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/436,717  
CC FILING DATE: 20-DEC-1993  
CC CLASSIFICATION: 530  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/JP92/00544  
CC FILING DATE: 24-APR-1992  
CC APPLICATION NUMBER: JP 4-32084  
CC FILING DATE: 19-FEB-1992  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: JP 3-95476  
CC FILING DATE: 25-APR-1991  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: WEGNER, Harold C.  
CC REGISTRATION NUMBER: 25,258  
CC REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (202)672-5300  
CC TELEFAX: (202)672-5399  
CC TELEX: 904136  
CC INFORMATION FOR SEQ ID NO: 100:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 135 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 135 AA: 14850 MW; 105384 CN;

Query Match 51.7%; Score 45; DB 2; Length 135;  
Best Local Similarity 55.6%; Pred. No. 2.07e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 73 FNGGTSYQ 81  
||| :||  
QY 62 FNGDEAYDQ 70

RESULT 12  
ID US-08-137-117D-100 STANDARD; PRT; 135 AA.  
XX  
AC xxxxxx  
XX  
DT  
DE  
XX

Sequence 100, Application US/08137117D  
Sequence 100, Application US/08137117D  
Patent No. 5795965  
GENERAL INFORMATION:  
CC APPLICANT: TSUCHIYA, Masayuki  
CC APPLICANT: SATO, Koh  
CC APPLICANT: BENDIG, Mary  
CC APPLICANT: JONES, Steven  
CC APPLICANT: SALDANHA, Jose  
CC TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
CC TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR  
CC NUMBER OF SEQUENCES: 158  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Foley & Lardner  
CC STREET: 3000 K Street, N.W., Suite 500  
CC CITY: Washington  
CC STATE: D.C.  
CC COUNTRY: USA  
CC ZIP: 20007-5109  
CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/137,117D  
CC FILING DATE: 20-DEC-1993  
CC CLASSIFICATION: 530  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/JP92/00544  
CC FILING DATE: 24-APR-1992  
CC APPLICATION NUMBER: JP 4-32084  
CC FILING DATE: 19-FEB-1992  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: JP 3-95476  
CC FILING DATE: 25-APR-1991  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: WEGNER, Harold C.  
CC REGISTRATION NUMBER: 25,258  
CC REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (202)672-5300  
CC TELEFAX: (202)672-5399  
CC TELEX: 904136  
CC INFORMATION FOR SEQ ID NO: 100:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 135 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 135 AA: 14850 MW; 105384 CN;

Query Match 51.7%; Score 45; DB 1; Length 135;  
Best Local Similarity 55.6%; Pred. No. 2.07e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 73 FNGGTSYQ 81  
||| :||  
QY 62 FNGDEAYDQ 70

RESULT 13  
ID US-08-137-117D-102 STANDARD; PRT; 135 AA.  
XX  
AC xxxxxx  
XX  
DT  
DE  
XX

Sequence 102, Application US/08137117D  
Sequence 102, Application US/08137117D  
Patent No. 5795965  
GENERAL INFORMATION:  
CC APPLICANT: TSUCHIYA, Masayuki  
CC APPLICANT: SATO, Koh  
CC APPLICANT: BENDIG, Mary  
CC APPLICANT: JONES, Steven  
CC APPLICANT: SALDANHA, Jose  
CC TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
CC TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR  
CC NUMBER OF SEQUENCES: 158  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Foley & Lardner  
CC STREET: 3000 K Street, N.W., Suite 500  
CC CITY: Washington  
CC STATE: D.C.  
CC COUNTRY: USA  
CC ZIP: 20007-5109  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/137,117D  
CC FILING DATE: 20-DEC-1993  
CC CLASSIFICATION: 530  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/JP92/00544  
CC FILING DATE: 24-APR-1992  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: JP 4-32084  
CC FILING DATE: 19-FEB-1992  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: JP 3-95476  
CC FILING DATE: 25-APR-1991  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: WEGNER, Harold C.  
CC REGISTRATION NUMBER: 25,258  
CC REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (202)672-5300  
CC TELEFAX: (202)672-5399  
CC TELEX: 904136  
CC INFORMATION FOR SEQ ID NO: 102:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 135 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: Protein  
CC SEQUENCE 135 AA; 14860 MW; 106394 CN;

Query Match 51.7%; Score 45; DB 1; Length 135;  
Best Local Similarity 55.6%; Pred. No. 2.07e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 73 FNGGTSYQ 81  
QY 62 FNGDEAYDQ 70

RESULT 14  
ID US-08-137-117D-112 STANDARD: PRT: 135 AA.

XX xxxxxx

Sequence 112, Application US/08137117D  
Sequence 112, Application US/08137117D  
Patent No. 5795965

GENERAL INFORMATION:  
APPLICANT: TSUCHIYA, Masayuki  
APPLICANT: SATO, Koh  
APPLICANT: BENDIG, Mary  
APPLICANT: JONES, Steven  
APPLICANT: SALDANHA, Jose  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR  
NUMBER OF SEQUENCES: 158  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,117D

CC FILING DATE: 20-DEC-1993  
CC CLASSIFICATION: 530  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: WO PCT/JP92/00544  
CC FILING DATE: 24-APR-1992  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: JP 4-32084  
CC FILING DATE: 19-FEB-1992  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: JP 3-95476  
CC FILING DATE: 25-APR-1991  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: WEGNER, Harold C.  
CC REGISTRATION NUMBER: 25,258  
CC REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (202)672-5300  
CC TELEFAX: (202)672-5399  
CC TELEX: 904136  
CC INFORMATION FOR SEQ ID NO: 112:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 135 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 135 AA; 14974 MW; 104899 CN;

Query Match 51.7%; Score 45; DB 1; Length 135;  
Best Local Similarity 55.6%; Pred. No. 2.07e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 73 FNGGTSYQ 81  
QY 62 FNGDEAYDQ 70

RESULT 15  
ID US-08-137-117D-27 STANDARD: PRT: 135 AA.

XX xxxxxx

Sequence 27, Application US/08137117D  
Sequence 27, Application US/08137117D  
Patent No. 5795965

GENERAL INFORMATION:  
APPLICANT: TSUCHIYA, Masayuki  
APPLICANT: SATO, Koh  
APPLICANT: BENDIG, Mary  
APPLICANT: JONES, Steven  
APPLICANT: SALDANHA, Jose  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR  
NUMBER OF SEQUENCES: 158  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,117D  
FILING DATE: 20-DEC-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: WO PCT/JP92/00544  
 CC FILING DATE: 24-APR-1992  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: JP 4-32084  
 CC FILING DATE: 19-FEB-1992  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: JP 3-95476  
 CC FILING DATE: 25-APR-1991  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: WEGNER, Harold C.  
 CC REGISTRATION NUMBER: 25,258  
 CC REFERENCE/DOCKET NUMBER: 53466/126/AOK  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (202)672-5300  
 CC TELEFAX: (202)672-5399  
 CC TELEX: 904136  
 CC INFORMATION FOR SEQ ID NO: 27:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 135 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 135 AA; 14701 MW; 106494 CN;

Query Match 51.7%; Score 45; DB 1; Length 135;  
 Best Local Similarity 55.8%; Pred. No. 2.07e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 73 FNGGTSYNQ 81  
 QY 62 FNGDEAYDQ 70

Search completed: Wed May 10 12:35:34 2000  
 Job time : 52 secs.

\*\*\*\*\*  
 M P S R L H  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (C) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed May 10 12:21:20 2000; MasPar time 227.71 Seconds  
 2.740 Million cell updates/sec  
 Tabular output not generated.

Title: >US-09-376-430-2  
 Description: (48-56) from US09376430A.pep (10 of 25)  
 Perfect Score: 58  
 Sequence: 1 ASKYSRTNL 9

Scoring table: PAM 150  
 Gap 11

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: spiremb112  
 1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
 5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
 9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
 13:sp-vertebrate 14:sp-virus

Statistics: Mean 21.817; Variance 23.713; scale 0.920

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	50	86.2	946	3	CHROMOSOME XII READING	1.36e+00
2	48	82.8	160	10	F15K9.23 PROTEIN.	4.03e+00
3	47	81.0	638	2	OMPLE (FRAGMENT).	6.86e+00
4	47	81.0	671	2	FUNCTION UNKNOWN.	6.86e+00
5	46	79.3	230	2	UBIQUINONE METHYLTRANS	1.16e+01
6	46	79.3	422	2	PUTATIVE MEMBRANE PROT	1.16e+01
7	46	79.3	523	11	CDC-LIKE PROTEIN (FRAG	1.16e+01
8	46	79.3	529	2	HYPOTHETICAL 60.9 KD P	1.16e+01
9	46	79.3	802	4	POMBE CDC5-RELATED PRO	1.16e+01
10	46	79.3	1152	2	2-ACYLGLYCEROPHOSPHOET	1.16e+01
11	45	77.6	496	10	GLUTAMATE DECARBOXYLAS	1.95e+01
12	45	77.6	496	10	GLUTAMATE DECARBOXYLAS	1.95e+01
13	45	77.6	496	10	GLUTAMATE DECARBOXYLAS	1.95e+01
14	45	77.6	536	2	CATALASE (EC 1.11.1.6)	1.95e+01
15	45	77.6	569	10	L-ASCORBATE OXIDASE PR	1.95e+01
16	45	77.6	595	2	HYPOTHETICAL 68.4 KD P	1.95e+01
17	45	77.6	1807	3	HYPOTHETICAL 229.9KD P	1.95e+01
18	45	77.6	1919	3	HYPOTHETICAL 217.7 KD	1.95e+01
19	45	75.9	172	5	ARP-LIKE PROTEIN.	3.23e+01
20	44	75.9	341	14	PPG1L, ORTHOLOGUE OF V	3.23e+01

ALIGNMENTS

RESULT ID	1	PRELIMINARY;	PRT;	946 AA.
AC	Q12369			
DT	Q12369			
DT	01-NOV-1996 (TREMELrel. 01, Created)			
DT	01-NOV-1996 (TREMELrel. 01, Last sequence update)			
DT	01-NOV-1998 (TREMELrel. 08, Last annotation update)			
DE	CHROMOSOME XII READING FRAME ORF YLL003W.			
DN	SFI1 OR L1373..			
OS	Saccharomyces cerevisiae (Baker's Yeast).			
OC	Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;			
OC	Saccharomycetales; Saccharomycetes.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MIOGA T., ZIMMERMANN F.K.;			
RL	Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	MIPS;			
RL	Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MA P., VAN DIJCK P., WINDERICKX J., THEVELEIN J.M.;			
RL	Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	STRAIN=FY23 /RD005 (DERIVED FROM S288C);			
RX	MEDLINE; 96405918.			
RA	MIOGA T., ZIMMERMANN F.K.;			
RT	"Sequence analysis of the <sub>1</sub> CEN12 region of Saccharomyces cerevisiae on			
RT	a 43.7 kb fragment of chromosome XII including an open reading frame			
RT	homologous to the human cystic fibrosis transmembrane conductance			
RT	regulator protein CFTR."			
RL	Yeast 12:693-708(1996).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RA	STRAIN=FY23 /RD005 (DERIVED FROM S288C);			
RX	MEDLINE; 91094833.			
RA	DAMAK F., BOYFAROCOTTE E., LE-ROSCOUET D., GUILBAUD R., JACQUET M.;			
RT	"SDC25, a CDC25-like gene which contains a RAS-activating domain and			
RT	is a dispensable gene of Saccharomyces cerevisiae."			
RL	Mol. Cell. Biol. 11:202-212(1991).			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RA	STRAIN=FY23 /RD005 (DERIVED FROM S288C);			

21	44	75.9	578	2	Q9X7P5	3.22e+01
22	44	75.9	789	14	Q9263	3.22e+01
23	44	75.9	899	5	Q19885	3.22e+01
24	44	75.9	1272	3	Q13756	3.22e+01
25	43	74.1	223	1	Q58970	5.33e+01
26	43	74.1	255	5	Q9XV62	5.33e+01
27	43	74.1	262	1	Q30290	5.33e+01
28	43	74.1	325	2	Q84076	5.33e+01
29	43	74.1	342	4	Q75796	5.33e+01
30	43	74.1	464	9	Q37899	5.33e+01
31	43	74.1	531	2	Q927F9	5.33e+01
32	43	74.1	535	3	Q13622	5.33e+01
33	43	74.1	592	10	Q23433	5.33e+01
34	43	74.1	615	2	Q86143	5.33e+01
35	43	74.1	903	14	Q69076	5.33e+01
36	43	74.1	904	14	Q69526	5.33e+01
37	43	74.1	1518	2	Q00600	5.33e+01
38	43	74.1	1792	6	Q46385	5.33e+01
39	42	72.4	242	5	Q20350	8.71e+01
40	42	72.4	277	2	Q32929	8.71e+01
41	42	72.4	369	1	Q58163	8.71e+01
42	42	72.4	405	3	Q99079	8.71e+01
43	42	72.4	467	3	Q74430	8.71e+01
44	42	72.4	740	5	Q9XUM0	8.71e+01
45	42	72.4	1904	9	Q38319	8.71e+01

PUTATIVE BCCT FAMILY T  
 COUNTERPART OF HSV-1 G  
 COSMID F28F5.  
 HYPOTHETICAL 146.2 KD  
 223AA LONG HYPOTHETICA  
 CHYMOTRYPSIN-LIKE SERI  
 5.33e+01  
 IRON-SULFUR CLUSTER BI  
 5.33e+01  
 PREDICTED OMP.  
 PROSTATE APOPTOSIS RES  
 5.33e+01  
 MAJOR TAIL PROTEIN GP2  
 5.33e+01  
 CT632 HYPOTHETICAL PRO  
 5.33e+01  
 ATP-DEPENDENT RNA HELI  
 5.33e+01  
 ECTP PROTEIN.  
 5.33e+01  
 GLYCOPROTEIN B.  
 5.33e+01  
 GLYCOPROTEIN B.  
 5.33e+01  
 GLUCOSYLTRANSFERASE I  
 5.33e+01  
 SUPERVILLIN.  
 SIMILAR TO MICROTUBULE  
 5.33e+01  
 HYPOTHETICAL 30.1 KD P  
 8.71e+01  
 369AA LONG HYPOTHETICA  
 8.71e+01  
 VINILOZOLIN RESISTANCE  
 8.71e+01  
 TRIGLYCERIDE LIPASE-CH  
 8.71e+01  
 W04A8.6 PROTEIN.  
 8.71e+01  
 ORF1904.  
 8.71e+01

RA MEDLINE; 93087480.  
 RA RIFMASTER T.L., VAUGHN G.P., WOOLFORD J.L. JR.;  
 RT "A putative ATP-dependent RNA helicase involved in Saccharomyces  
 RT cerevisiae ribosome assembly.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:11131-11135(1992).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FY23 /RD005 (DERIVED FROM S288C);  
 RX MEDLINE; 94375516.  
 RA BURGESS S.M., DELANNOY M., JENSEN R.E.;  
 RT "MM1 encodes a mitochondrial outer membrane protein essential for  
 RT establishing and maintaining the structure of yeast mitochondria.";  
 RL J. Cell Biol. 126:1375-1391(1994).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FY23 /RD005 (DERIVED FROM S288C);  
 RX MEDLINE; 95348179.  
 RA GAMMIE A.E., KURIHARA L.J., VALLEE R.B., ROSE M.D.;  
 RT "DNM1, a dynamin-related gene, participates in endosomal trafficking  
 RT in yeast.";  
 RL J. Cell Biol. 130:553-566(1995).  
 DR EMBL; 773108; CAA97446.1; -.  
 DR EMBL; X95569; CAA64815.1; -.  
 DR EMBL; X91488; CAA62767.1; -.  
 SQ SEQUENCE 946 AA; 112578 MW; BC59DC8C CRC32;

Query Match 86.2%; Score 50; DB 3; Length 946;  
 Best Local Similarity 88.9%; Pred. No. 1.36e+00;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 476 ASDYSRTNL 484  
 QY 48 ASKYSRTNL 56

RESULT 2  
 ID Q9VVR8 PRELIMINARY; PRT; 160 AA.  
 AC Q9VVR8;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE F15K9.23 PROTEIN.  
 GN F15K9.23.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA VYSOTSKAIA V.S., SCHWARTZ J.R., TORIUMI M., YU G., LI J., LIU S.,  
 RA KREMENETSKAIA I., LUKOS J., ARAUJO R., BUEHLER E., CONWAY A.B.,  
 RA DEWER K., FENG J., KIM C., LI Y., SHINN P., SUN H., DAVIS R.W.,  
 RA ECKER J.R., FEDERSPEL N.A., THEOLOGIS A.;  
 RT "Arabidopsis thaliana chromosome 1 BAC F15K9 sequence.";  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA THEOLOGIS A.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA THEOLOGIS A.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA THEOLOGIS A.;  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC005278; AAC72117.1; -.

SQ SEQUENCE 160 AA; 18464 MW; FC1D4959 CRC32;  
 Query Match 82.8%; Score 48; DB 10; Length 160;  
 Best Local Similarity 66.7%; Pred. No. 4.03e+00;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 12 SSYSRNL 20  
 QY 48 ASKYSRTNL 56

RESULT 3  
 ID Q46519 PRELIMINARY; PRT; 638 AA.  
 AC Q46519;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE OMPL1 (FRAGMENT).  
 GN OMPL1.  
 OS Bacteroides nodosus (Dichelobacter nodosus).  
 OC Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;  
 OC Dichelobacter.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VCS1001 (A198);  
 RX MEDLINE; 96020672.  
 RA MOSES E.K., GOOD R.T., SINISTAJ M., BILLINGTON S.J., LANGFORD C.J.,  
 RA ROOD J.I.;  
 RT "A multiple site-specific DNA-inversion model for the control of Omp1  
 RT phase and antigenic variation in Dichelobacter nodosus.";  
 RL Mol. Microbiol. 17:183-196(1995).  
 DR EMBL; U02462; AAB12361.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 638 AA; 69006 MW; 44F37A4C CRC32;

Query Match 81.0%; Score 47; DB 2; Length 638;  
 Best Local Similarity 75.0%; Pred. No. 6.86e+00;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Db 229 AKYERTNL 236  
 QY 49 SKYSRTNL 56

RESULT 4  
 ID Q45597 PRELIMINARY; PRT; 671 AA.  
 AC Q45597;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE FUNCTION UNKNOWN.  
 GN YDE OR FBP.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RX MEDLINE; 96051385.  
 RA OGASAWARA N., NARAI S., YOSHIKAWA H.;  
 RT "Systematic sequencing of the 180 kilobase region of the Bacillus  
 RT subtilis chromosome containing the replication origin.";  
 RL DNA Res. 1:11-14(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168;  
 RX MEDLINE; 94171085.  
 RA ZHANG J., ARONSON A.I.;  
 RT "A Bacillus subtilis bglA gene encoding phospho-beta-glucosidase is  
 RT inducible and closely linked to a NADH dehydrogenase-encoding gene.";  
 RL Gene 140:85-90(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.



RC STRAIN-168;  
RX MEDLINE; 94156824.  
RA CALOGERO S., GARDAN R., GLASER P., SCHWEIZER J., RAPOPORT G.,  
RA DEBARBOUILLE M.;  
RT "RocR, a novel regulatory protein controlling arginine utilization in  
RT Bacillus subtilis, belongs to the NtrC/NifA family of transcriptional  
RT activators";  
RL J. Bacteriol. 176:1234-1241(1994).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RX MEDLINE; 94236234.  
RA HARTFORD O.M., DOWDS B.C.;  
RT "Isolation and characterization of a hydrogen peroxide resistant  
RT mutant of Bacillus subtilis";  
RL Microbiology 140:297-304(1994).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RX MEDLINE; 96093926.  
RA YOSHIDA K., SEKI S., FUJIMURA M., MIWA Y., FUJITA Y.;  
RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis  
RT genome between the gnt and iol operons";  
RL DNA Res. 2:61-69(1995).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RX MEDLINE; 95311309.  
RA GARDAN R., RAPOPORT G., DEBARBOUILLE M.;  
RT "Expression of the rocDEF operon involved in arginine catabolism in  
RT Bacillus subtilis";  
RL J. Mol. Biol. 249:843-856(1995).  
RN [7]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RA KASAHARA Y., NAKAI S., YOSHIKAWA H., OGASAWARA N.;  
RT Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
RN [8]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RX MEDLINE; 98044033.  
RA KUNT E., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,  
RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,  
RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,  
RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,  
RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMERSON P.T.,  
RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,  
RA FRITZ C., FUJITA M., FUJITA Y., FUNA S., GALIZZI A., GALLERON N.,  
RA GHIM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,  
RA GUISEPPI G., GUY B.J., HAGA K., HATECH J., HARWOOD C.R., HENAUT A.,  
RA HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,  
RA JORTS B., KARAMATA D., KASAHARA Y., KLAER-BLANCHARD M., KLEIN C.,  
RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M.,  
RA KURITA K., LAPIDUS A., LARDINOS S., LAUBER J., LAZAREVIC V.,  
RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,  
RA MEDINA N., MELIADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,  
RA NEONE D., O'REILLY M., OGAWA K., OGAWARA A., OUDEGA B., PARK S.H.,  
RA PARRO V., POHL T.M., PORTEILLE D., PORWOLLIK S., PRESOTTI A.M.,  
RA PRESCAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,  
RA RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADAIE Y.,  
RA SATO T., SCANLAN E., SCHLEICHER S., SCHROETER R., SCOFFONE F.,  
RA SEKIGUCHI J., SEKOWSKA A., SERO S.J., SERROR P., SHIN B.S., SOLDI B.,  
RA SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,  
RA TAKEUCHI M., TAMAKOSHI A., TANAKA T., TERESTRA P., TOGNONI A.,  
RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,  
RA VIARI A., WAMBUIT R., WEDLER E., WEDLER H., WEITZNEGGER T.,  
RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,  
RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.;  
RT "The complete genome sequence of the gram-positive bacterium Bacillus  
RT subtilis";  
RL Nature 390:249-256(1997).  
RN [9]

RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RA KUNT E., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; D78193; BAA11277.1; -;  
DR EMBL; Z91124; CAB16056.1; -;  
KW Hydrolase.  
SQ SEQUENCE 671 AA; 77978 MW; 1EBB6B4D CRC32;  
  
Query Match 81.0%; Score 47; DB 2; Length 671;  
Best Local Similarity 55.6%; Pred. No. 6.86e+00;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
  
Db 154 SSKYTRSKL 162  
QY 48 ASKYSRTNL 56  
  
RESULT 5 PRELIMINARY; PRT; 230 AA.  
ID Q92837;  
AC Q92837;  
DT 01-MAY-1999 (TREMELrel. 10, Created)  
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)  
DE UBIQUINONE METHYLTRANSFERASE.  
GN UBIE.  
OS Chlamydia pneumoniae.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiaophila.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CW1029;  
RA KALMAN S., MITCHELL W., MARATHE R., LAMMEL C., FAN J., OLINGER L.,  
RA GRIMWOOD J., DAVIS R.W., STEPHENS R.S.;  
RT "Comparative Genomes of Chlamydia pneumoniae and C. trachomatis";  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE001636; AAD18655.1; -;  
KW Transferase.  
SQ SEQUENCE 230 AA; 25637 MW; 4F18DD4E CRC32;  
  
Query Match 79.3%; Score 46; DB 2; Length 230;  
Best Local Similarity 87.5%; Pred. No. 1.16e+01;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Db 18 ASKYDRTN 25  
QY 48 ASKYSRTN 55  
  
RESULT 6 PRELIMINARY; PRT; 422 AA.  
ID O50548;  
AC O50548;  
DT 01-JUN-1998 (TREMELrel. 06, Created)  
DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)  
DT 01-NOV-1999 (TREMELrel. 12, Last annotation update)  
DE PUTATIVE MEMBRANE PROTEIN.  
GN ESTX OR TM1021.  
OS Thermotoga maritima.  
OC Bacteria; Thermotogales; Thermotoga.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MSB8 (DSM3109);  
RA LIEBL W., GABELSBERGER J.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 99287316.  
RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,  
RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,  
RA STEWART A.M., COTTON M.D., MALEK J.A., LINHER K.D., GARRETT M.M.,  
RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,  
RA SMITH H.O., VENTER J.C., FRASER C.M.;  
RT "Evidence for lateral gene transfer between Archaea and bacteria from

```

RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,
RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,
RA MCDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,
RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,
RA SMITH H.O., VENTER J.C., FRASER C.M.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ001694; CA04933.1; -
DR EMBL: AE001764; AAD36098.1; -
DR TIGR: TM1021; -
DR PFAM: PF00083; sugar_tr; 1.
SQ SEQUENCE 422 AA; 46050 MW; A255D191 CRC32;

Query Match 79.3%; Score 46; DB 2; Length 422;
Best Local Similarity 66.7%; Pred. No. 1.16e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 60 ADYRSKNL 68
QY 48 ASKYSRTNL 56

RESULT 7
ID O08837 PRELIMINARY; PRT; 523 AA.
AC O08837;
DT 01-JUL-1997 (TremBLrel. 04, Created)
DT 01-JUL-1997 (TremBLrel. 04, Last sequence update)
DT 01-AUG-1998 (TremBLrel. 07, Last annotation update)
DE CDC-LIKE PROTEIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NOBLE;
RA TOO C.K.L.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF000578; AAD05365.1; -
FT NON_TER 1
SQ SEQUENCE 523 AA; 59507 MW; 81200757 CRC32;

Query Match 79.3%; Score 46; DB 11; Length 523;
Best Local Similarity 62.5%; Pred. No. 1.16e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 396 SRYTRANL 403
QY 49 SKYSRTNL 56

RESULT 8
ID O84637 PRELIMINARY; PRT; 529 AA.
AC O84637;
DT 01-NOV-1998 (TremBLrel. 08, Created)
DT 01-NOV-1998 (TremBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TremBLrel. 08, Last annotation update)
DE HYPOTHETICAL 60.9 KD PROTEIN.
CN CT632.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaeae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D/UW-3/CX;
RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L.,
RA MITCHELL W.P., OLINGER L., TATSOV R.L., ZHAO Q., KOONIN E.V.,
RA DAVIS R.W.;
RL "Genome Sequence of an Obligate Intracellular Pathogen of Humans:
RT Chlamydia trachomatis."
RL Science 0:0-0(1998).

```

---

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-D/UW-3/CX;
RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L.,
RA MITCHELL W.P., OLINGER L., TATSOV R.L., ZHAO Q., KOONIN E.V.,
RA DAVIS R.W.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE001334; AAC68236.1; -
KW Hypothetical protein
SQ SEQUENCE 529 AA; 60915 MW; D6328EC8 CRC32;

Query Match 79.3%; Score 46; DB 2; Length 529;
Best Local Similarity 75.0%; Pred. No. 1.16e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 43 SKYSRTNL 50
QY 49 SKYSRTNL 56

RESULT 9
ID O9974 PRELIMINARY; PRT; 802 AA.
AC Q9974;
DT 01-MAY-1997 (TremBLrel. 03, Created)
DT 01-MAY-1997 (TremBLrel. 03, Last sequence update)
DT 01-NOV-1999 (TremBLrel. 12, Last annotation update)
DE POMBE CDC5-RELATED PROTEIN, KIAA0432.
GN PCDC5RP OR KIAA0432.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97190317.
RA BERNSTEIN H.S., COUGHLIN S.R.;
RT "Pombe Cdc5-related protein. A putative human transcription factor
RT implicated in mitogen-activated signaling."
RL J. Biol. Chem. 272:5833-5837(1997).
RN [2]
RP SEQUENCE OF 54-802 FROM N.A.
RC TISSUE-BRAIN;
RA ISHIKAWA K., NAGASE T., NAKAJIMA D., SEKI N., OHIRA M., MIYAJIMA N.,
RA TANAKA A., KOTANI H., NOMURA N., OHARA O.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U86753; AAB61210.1; -
DR EMBL: AB007892; BAA24862.1; -
DR HSP; P06876; 1MBH
DR PFAM: PF00249; myb_DNA-binding; 2.
SQ SEQUENCE 802 AA; 92250 MW; 04CB9B56 CRC32;

Query Match 79.3%; Score 46; DB 4; Length 802;
Best Local Similarity 62.5%; Pred. No. 1.16e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 675 SRYTRANL 682
QY 49 SKYSRTNL 56

RESULT 10
ID O9ZCU1 PRELIMINARY; PRT; 1152 AA.
AC Q9ZCU1;
DT 01-MAY-1999 (TremBLrel. 10, Created)
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)
DT 01-NOV-1999 (TremBLrel. 12, Last annotation update)
DE 2-ACVLYGLYCEROPHOSPHOETHANOLAMINE ACYLTRANSFERASE (AAS).
GN RP620.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MADRID E;

```

RX MEDLINE; 99039499.  
RA ANDERSSON S.G.E., ZOMORODIPOUR A., ANDERSSON J.O.,  
RA SICHERITZ-PONTEN T., ALSMARK U.C.M., PODOWSKI R.M., NAESLUND A.K.,  
RA ERIKSSON A.S., WINKLER H.H., KURLAND C.G.;  
RT "The genome sequence of Rickettsia prowasekii and the origin of  
RT mitochondria";  
RL Nature 396:133-140(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MADRID E;  
RA ANDERSSON S.G.E.;  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ235272; CAA15063.1; -;  
DR PROSITE; PS00455; AMP-BINDING; 1.  
SQ SEQUENCE 1152 AA; 130356 MW; A61294CC CRC32;

Query Match 79.3%; Score 46; DB 2; Length 1152;  
Best Local Similarity 66.7%; Pred. No. 1.16e+01;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 77 ADKYERANL 85  
QY 48 ASKYSRTNL 56

RESULT 11 PRELIMINARY; PRT; 496 AA.  
ID O81101;  
AC O81101;  
DT 01-NOV-1998 (TREMREL. 08, Created)  
DT 01-NOV-1998 (TREMREL. 08, Last sequence update)  
DT 01-NOV-1999 (TREMREL. 12, Last annotation update)  
DE GLUTAMATE DECARBOXYLASE ISOZYME 2.  
GN NTGAD2.  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
OC Nicotiana.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 98302498.  
RA YUN S.J., OH S.H.;  
RT "Cloning and characterization of a tobacco cDNA encoding  
RT calcium/calmodulin-dependent glutamate decarboxylase";  
RL Mol. Cells 8:125-129(1998).  
DR EMBL; AF020424; AAC39483.1; -;  
DR MENDEL; 31754; Nicta:1402:31754.  
DR PFAM; PF00282; PYRIDOXAL\_DEC; 1.  
SQ SEQUENCE 496 AA; 55931 MW; 2C268A0B CRC32;

Query Match 77.6%; Score 45; DB 10; Length 496;  
Best Local Similarity 66.7%; Pred. No. 1.95e+01;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 19 ASRYVRTSL 27  
QY 48 ASKYSRTNL 56

RESULT 12 PRELIMINARY; PRT; 496 AA.  
ID P93369;  
AC P93369;  
DT 01-MAY-1997 (TREMREL. 03, Created)  
DT 01-MAY-1997 (TREMREL. 03, Last sequence update)  
DT 01-NOV-1999 (TREMREL. 12, Last annotation update)  
DE GLUTAMATE DECARBOXYLASE.  
GN NTGAD1.  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
OC Nicotiana.  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-XANTHI;  
RA DHARMASIRI M.A.N., LU Y.T., HARRINGTON H.M.;  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U54774; AAB40608.1; -;  
DR MENDEL; 9386; Nicta:1402:9386.  
DR PFAM; PF00282; PYRIDOXAL\_DEC; 1.  
SQ SEQUENCE 496 AA; 56035 MW; 726D3A56 CRC32;  
Query Match 77.6%; Score 45; DB 10; Length 496;  
Best Local Similarity 66.7%; Pred. No. 1.95e+01;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 19 ASRYVRTSL 27  
QY 48 ASKYSRTNL 56

RESULT 13 PRELIMINARY; PRT; 496 AA.  
ID O81102;  
AC O81102;  
DT 01-NOV-1998 (TREMREL. 08, Created)  
DT 01-NOV-1998 (TREMREL. 08, Last sequence update)  
DT 01-NOV-1999 (TREMREL. 12, Last annotation update)  
DE GLUTAMATE DECARBOXYLASE ISOZYME 1.  
GN NTGAD1.  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;  
OC Nicotiana.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA YUN S.J., OH S.H.;  
RT "Cloning and characterization of a tobacco cDNA encoding  
RT calcium/calmodulin-dependent glutamate decarboxylase";  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF020425; AAC24195.1; -;  
DR MENDEL; 31755; Nicta:1402:31755.  
DR PFAM; PF00282; PYRIDOXAL\_DEC; 1.  
SQ SEQUENCE 496 AA; 55963 MW; 7BA418F5 CRC32;

Query Match 77.6%; Score 45; DB 10; Length 496;  
Best Local Similarity 66.7%; Pred. No. 1.95e+01;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 19 ASRYVRTSL 27  
QY 48 ASKYSRTNL 56

RESULT 14 PRELIMINARY; PRT; 536 AA.  
ID Q59337;  
AC Q59337;  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-NOV-1999 (TREMREL. 12, Last annotation update)  
DE CATALASE (EC 1.11.1.6).  
GN KATA.  
OS Deinococcus radiodurans.  
OC Bacteria; Thermus/Deinococcus group; Deinococcus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-KR1;  
RA NARUMI I., WATANABE H., HOSAIN A., TANAKA A., KITAYAMA S.;  
RT "Molecular cloning and nucleotide sequence of radiation-inducible  
RT catalase gene from radioresistant bacterium, Deinococcus  
RT radiodurans";  
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; D63899; BAA0937.1; -;  
DR HSSP; P00432; 4BLC.  
DR PFAM; PF00199; catalase; 1.  
DR PRINTS; PR00067; CATALASE.  
RN [1]

KW Oxidoreductase; Peroxidase.  
SQ SEQUENCE 536 AA; 60622 MW; 2A1BADB7 CRC32;

Query Match 77.6%; Score 45; DB 2; Length 536;  
Best Local Similarity 62.5%; Pred.No.1.95e+01;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 103 SKYTRAKL 110  
||||:|  
QY 49 SKYSRTNL 56

RESULT 15  
ID O24093 PRELIMINARY; PRT; 569 AA.  
AC O24093;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
DE L-ASCORBATE OXIDASE PRECURSOR.  
GN MTN23.  
OS Medicago truncatula (Barrel medic).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
OC core eudicots; Rosidae; eutosids I; Fabales; Fabaceae; Papilionoideae;  
OC Medicago.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. JEMALONG J5; TISSUE=ROOT NODULE;  
RX MEDLINE; 96212994.  
RA GAVAS P.; DE CARVALHO NIEBEL F.; LESCURE N.; CULLIMORE J.;  
RT "Use of a subtractive hybridization approach to identify new Medicago  
truncatula genes induced during root nodule development.";  
RL Mol. Plant Microbe Interact. 9:233-242(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. JEMALONG J5; TISSUE=ROOT NODULE;  
RA GAVAS P.;  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Y15295; CAA75577.1; -.  
DR HSSP; P37064; IASP.  
DR MENDEL; 27077; Medtr;2501;27077.  
DR PROSITE; PS00079; MULTICOPPER\_OXIDASE1; 1.  
DR PROSITE; PS00080; MULTICOPPER\_OXIDASE2; 1.  
DR PEAR; PF00394; Cu-oxidase; 2.  
KW Signal.  
FT SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 569 POTENTIAL.  
SQ SEQUENCE 569 AA; 63973 MW; 7E79B9AD CRC32;

Query Match 77.6%; Score 45; DB 10; Length 569;  
Best Local Similarity 77.8%; Pred.No.1.95e+01;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 205 ASKYGSTNL 213  
||||:|  
QY 48 ASKYSRTNL 56

Search completed: Wed May 10 12:25:19 2000  
Job time : 239 secs.

\*\*\*\*\*  
M P S R L H  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed May 10 12:34:16 2000; MasPar time 2.93 Seconds  
Tabular output not generated. 96.912 Million cell updates/sec

Title: >US-09-376-430-2  
Description: (62-73) from US09376430A.pap (11 of 25)  
Perfect Score: 87  
Sequence: 1 FNGDEAYDOCTN 12

Scoring table: PAM 150  
Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq35  
1:geneseqp

Statistics: Mean 17.278; Variance 51.238; scale 0.337

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	55	63.2	663	1	W41111 Chicken matrix metallo	3.25e+01
2	55	63.2	663	1	W41227 Chicken matrix metallo	3.25e+01
3	53	60.9	170	1	H. pylori strain J166	5.40e+01
4	51	58.6	429	1	Human matrix metallopr	8.92e+01
5	51	58.6	631	1	Complete type IV colla	8.92e+01
6	51	58.6	631	1	Human mature matrix me	8.92e+01
7	51	58.6	631	1	Human type IV collagen	8.92e+01
8	51	58.6	631	1	Sequence of human type	8.92e+01
9	51	58.6	660	1	Type IV collagenase CD	8.92e+01
10	48	55.2	73	1	Hirudin (HV-3) CRGMP	1.87e+02
11	48	55.2	73	1	Hirudin (HV-3) CRGMP	1.87e+02
12	48	55.2	178	1	H. pylori strain 60190	1.87e+02
13	48	55.2	178	1	Plasmodium PRP100 ICA	1.87e+02
14	48	55.2	241	1	Rhymiculate phosphonyd	1.87e+02
15	47	54.0	411	1	Sequence of the envelo	2.38e+02
16	47	54.0	521	1	Sequence encoded by al	2.38e+02
17	47	54.0	671	1	Wheat soluble starch s	2.38e+02
18	47	54.0	922	1	Trimmed enzyme protein	2.38e+02
19	47	54.0	1230	1	Potato tuber soluble s	2.38e+02
20	46	52.9	339	1	Protein activated lipa	3.03e+02
21	46	52.9	430	1	Human MAD-related prot	3.03e+02
22	46	52.9	466	1	Xenopus alpha-signalin	3.03e+02
23	46	52.9	467	1	Human MAD-related prot	3.03e+02

24	46	52.9	632	1	W99016 Human matrilin-3.	3.03e+02
25	46	52.9	2670	1	R88125 Rat IP3 receptor.	3.03e+02
26	46	52.9	2670	1	R88126 Human IP3 receptor.	3.03e+02
27	45	51.7	107	1	W57219 Cytochrome c3 protein	3.85e+02
28	45	51.7	114	1	R21271 Murine VH group 1 chain	3.85e+02
29	45	51.7	116	1	W22418 Reshaped human AUK12-2	3.85e+02
30	45	51.7	197	1	W56222 Smad1 protein C-termin	3.85e+02
31	45	51.7	198	1	W56225 Smad5 protein C-termin	3.85e+02
32	45	51.7	199	1	W56244 Smad5 protein fragment	3.85e+02
33	45	51.7	236	1	R45442 Sequence of the single	3.85e+02
34	45	51.7	256	1	Y10987 H. pylori ORF 066e3070	3.85e+02
35	45	51.7	413	1	W41058 Phytoene synthase from	3.85e+02
36	45	51.7	428	1	W45963 Amino acid sequence of	3.85e+02
37	45	51.7	429	1	W45964 Amino acid sequence of	3.85e+02
38	45	51.7	464	1	W18094 Xenopus alpha-signalin	3.85e+02
39	45	51.7	649	1	R47504 Pyrococcus furiosus al	3.85e+02
40	45	51.7	748	1	R24396 Prod. of the S gene of	3.85e+02
41	45	51.7	749	1	R42465 Feline infectious peri	3.85e+02
42	45	51.7	1005	1	W93955 Human 53BP2 protein.	3.85e+02
43	45	51.7	1454	1	R42467 Feline infectious peri	3.85e+02
44	45	51.7	1454	1	R24397 Prod. of the S gene of	3.85e+02
45	45	51.7	1594	1	P81183 Sequence of the peplo	3.85e+02

ALIGNMENTS

RESULT 1  
ID W41111 standard; Protein; 663 AA.

AC W41111;  
DT 08-JUN-1998 (first entry)  
DE Chicken matrix metalloproteinase-2.  
KW Matrix metalloproteinase-2; MMP-2; chMMP-2; chicken;  
KW Angiogenesis; inhibitor; antagonist; integrin alpha-v beta-3;  
KW vitronectin receptor; rheumatoid arthritis; tumour; metastasis;  
KW diabetic retinopathy; macular degeneration; restenosis; therapy.  
OS Gallus sp.  
FH Key Location/Qualifiers  
FT Peptide 1..26  
FT /label= Sig\_peptide  
PN WO9745137-A1.  
PD 04-DEC-1997.  
PR 30-MAY-1997; U09158.  
PR 31-MAY-1996; US-018733.  
PR 31-MAY-1996; US-015969.  
PA (SCRI ) SCRIPPS RES INST.  
PI Brooks P, Cheres DA;  
DR WPI: 98-032334/03.  
DR N-PSDB; V03995.  
PT Packaging material containing polypeptide antagonist of alphav,  
PT beta3 integrin - used for inhibition of angiogenesis, and for  
PT treating tumours, inflammation, eye diseases etc.  
PS Disclosure; Page 163-167; 234pp; English.  
CC This protein sequence comprises chicken matrix metalloproteinase-2  
CC (chMMP-2). The invention relates to the discovery that angiogenesis  
CC is mediated by the specific vitronectin receptor alpha-v beta-3,  
CC and that inhibition of alpha-v beta-3 function inhibits  
CC angiogenesis. Claimed antagonists of alpha-v beta-3 include  
CC C-terminal fragments (see W41083-94) of human or chicken MMP-2. An  
CC PCR amplification of the chMMP-2 coding region, cloning into e.g.  
CC pGEX-3X, and expression in E. coli as a fusion protein with  
CC glutathione-S-transferase. The antagonists can be used to inhibit  
CC angiogenesis in inflamed tissue (for treatment of arthritis or  
CC rheumatoid arthritis), in solid tumours or metastases (particularly  
CC to induce regression or inhibit tumour growth), and in ocular  
CC disorders such as diabetic retinopathy and macular degeneration, as  
CC well as to treat restenosis (all claimed).  
SQ Sequence 663 AA;

Query Match 63.2%; Score 55; DB 1; Length 663;  
Best Local Similarity 63.6%; Pred. No. 3.25e+01;  
Matches 7; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Db 294 FQG-QSYDQCT 303  
 QY 62 FNGDEAYDQCT 72

## RESULT 2

ID W41227 standard; protein: 663 AA.  
 AC W41227;  
 DT 09-JUN-1998 (first entry)  
 DE Chicken matrix metalloproteinase-2 (MMP-2) protein sequence.  
 KW Matrix metalloproteinase-2; MMP-2; alpha-v-beta-5 antagonist; treatment;  
 KW vitronectin receptor; inhibition; angiogenesis; integrin; tumour growth;  
 KW restenosis; neovascularisation.  
 OS Gallus sp.  
 PN WO9745447-A1.  
 PD 04-DEC-1997.  
 PF 30-MAY-1997; U09099.  
 PR 31-MAY-1996; US-018733.  
 PR 31-MAY-1996; US-015869.  
 PA (Scripps) SCRIPPS RES INST.  
 PI Brooks P, Cheresh DA, Friedlander M;  
 DR WPI; 98-041758/04.  
 PT Packaging material containing polypeptide antagonist of alphav,  
 PT beta5 integrin - used for inhibition of angiogenesis, and for  
 PT treating tumours, inflammation, eye diseases etc.  
 PS Disclosure: Fig 15A-B; 117pp; English.  
 CC The present sequence represents the chicken matrix metalloproteinase-2  
 CC (MMP-2) protein sequence. Fragments of this protein (W41234-39) are  
 CC able to act as alpha-v-beta-5 antagonists. Alpha-v-beta-5 is a  
 CC vitronectin receptor. Inhibitors of alpha-v-beta-5 can inhibit  
 CC angiogenesis. The specification describes a novel labelled package that  
 CC contains an inhibitor of angiogenesis i.e. an alpha-v-beta-5 antagonising  
 CC polypeptide that binds to integrin alpha-v-beta-5 and includes a part of  
 CC the C-terminal domain of MMP. The antagonists are used to inhibit  
 CC angiogenesis in inflamed tissue, in solid tumours or metastases, and in  
 CC a wide range of ocular disorders (e.g. diabetic or other forms of  
 CC retinopathy, neovascular glaucoma, or corneal transplants). They are  
 CC particularly used to induce regression or to inhibit growth of tumours.  
 CC The alpha-v-beta-5 antagonists can also be used to treat restenosis  
 CC caused by migration of smooth muscle cells following angioplasty and to  
 CC reduce blood supply to selected tissues. The antagonists particularly  
 CC inhibit neovascularisation where this is induced by cytokines,  
 CC e.g. transforming growth factor alpha, epidermal growth factor or  
 CC especially vascular endothelial growth factor.  
 CC Sequence 663 AA;

Query Match 63.2%; Score 55; DB 1; Length 663;  
 Best Local Similarity 63.6%; Pred. No. 3.25e-01;  
 Matches 7; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Db 294 FQG-QSYDQCT 303  
 QY 62 FNGDEAYDQCT 72

## RESULT 3

ID W63805 standard; Protein: 170 AA.  
 AC W63805;  
 DT 01-OCT-1998 (first entry)  
 DE H. pylori strain J166 IceA 1 protein fragment.  
 KW IceA; immunoassay; detection; ulcerogenic; gastric carcinoma; treatment;  
 KW peptic ulcer; immunisation; vaccine; protection.  
 OS Helicobacter pylori.  
 PN WO9743901-A1.  
 PD 27-NOV-1997.  
 PF 20-MAY-1997; U08558.  
 PR 20-MAY-1996; US-650528.  
 PA (UVA-) UNIV VANDERBILT.  
 PI Blaser MJ, Miller GG, Peek RM, Thompson SA;  
 DR WPI; 98-286350/25.  
 PT New Helicobacter pylori proteins - induced by contact with  
 PT epithelium and related DNA, are associated with ulcer formation,  
 PT useful in diagnosis and immunisation

PS Claim 5; Page 88: 107pp; English.  
 CC This sequence represents an IceA 1 variant protein fragment from  
 CC Helicobacter pylori strain J166. This protein or its fragments, are used  
 CC in standard immunoassays to detect H. pylori-specific antibodies,  
 CC particularly for diagnosis, especially antibodies characteristic of  
 CC IceA-positive strains which are ulcerogenic. Detecting presence of  
 CC IceA-positive strains also allows the risk of developing gastric  
 CC carcinoma to be assessed. Ligands, particularly antibodies, that  
 CC recognise IceA proteins are used to treat peptic ulcers, while  
 CC immunisation with IceA-negative H. pylori is used to protect against  
 CC infection (and its consequences such as ulcers, gastritis and gastric  
 CC cancer). Immunogenic IceA fragments, or the nucleic acid encoding them,  
 CC can also be used for vaccination. Antibodies (Ab) raised against IceA can  
 CC be used therapeutically or to screen other strains for homologous  
 CC proteins. Expression of IceA is strongly correlated with ulceration, so  
 CC detecting IceA allows differentiation between ulcerogenic and  
 CC non-ulcerogenic strains.  
 CC Sequence 170 AA;

Query Match 60.9%; Score 53; DB 1; Length 170;  
 Best Local Similarity 50.0%; Pred. No. 5.40e+01;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 127 YAGEAAYDAC 136  
 QY 62 FNGDEAYDQCT 71

## RESULT 4

ID W41112 standard; Protein: 429 AA.  
 AC W41112;  
 DT 08-JUN-1998 (first entry)  
 DE Human matrix metalloproteinase hMMP-2 (aa203-631).  
 KW Matrix metalloproteinase; MMP-2; hMMP-2; human; angiogenesis;  
 KW inhibitor; antagonist; integrin alpha-v beta-3;  
 KW vitronectin receptor; rheumatoid arthritis; tumour; metastasis;  
 KW diabetic retinopathy; macular degeneration; restenosis; therapy.  
 OS Homo sapiens.  
 PN WO9745137-A1.  
 PD 04-DEC-1997.  
 PF 30-MAY-1997; U09158.  
 PR 31-MAY-1996; US-018733.  
 PR 31-MAY-1996; US-015869.  
 PA (Scripps) SCRIPPS RES INST.  
 PI Brooks P, Cheresh DA;  
 DR WPI; 98-032334/03.  
 PT Packaging material containing polypeptide antagonist of alphav,  
 PT beta3 integrin - used for inhibition of angiogenesis, and for  
 PT treating tumours, inflammation, eye diseases etc.  
 PS Example 4; Page 177-179; 234pp; English.  
 CC This polypeptide comprises amino acid residues 203-631 of human  
 CC mature matrix metalloproteinase 2 (hMMP-2). It was produced by  
 CC recombinant methods involving PCR amplification (see V12509) of  
 CC hMMP-2 coding sequence and cloning into e.g. pGEX-lambda vector  
 CC for expression in E. coli as a glutathione-S-transferase fusion  
 CC protein. The invention relates to the discovery that angiogenesis  
 CC is mediated by the specific vitronectin receptor alpha-v beta-3, and  
 CC that inhibition of alpha-v beta-3 function inhibits angiogenesis.  
 CC Claimed antagonists of alpha-v beta-3 comprise C-terminal fragments  
 CC (see W41083-94) of human or chicken MMP-2, fusion polypeptides,  
 CC cyclic or linear polypeptides (see also W41098-110), derivatised  
 CC polypeptides, a monoclonal antibody or organic mimetic compound.  
 CC The antagonists are used to inhibit angiogenesis in: inflamed  
 CC tissue for treatment of arthritis or rheumatoid arthritis; solid  
 CC tumours or metastases, particularly to induce tumour regression or  
 CC inhibit growth of tumours; and in ocular disorders such as diabetic  
 CC retinopathy or macular degeneration (all claimed). They can also  
 CC be used to treat restenosis caused by migration of smooth muscle  
 CC cells following angioplasty and to reduce blood supply to selected  
 CC tissues (claimed). The new antagonists are highly selective for  
 CC angiogenesis. Only new blood vessels express alpha-v beta-3, so  
 CC mature vessels are unaffected, and the antagonists should be of low  
 CC toxicity.



```

SQ Sequence 631 AA;
Query Match 58.6%; Score 51; DB 1: Length 631;
Best Local Similarity 58.3%; Pred. No. 8.92e+01;
Matches 7; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Db 210 FNGKE-YNSTD 220
    ||| | | :
Qy 62 FNGDEAYDQCTN 73  ↓

RESULT 9
ID R06420 standard; protein: 660 AA.
AC R06420;
DT 13-DEC-1990 (first entry)
DE Type IV collagenase cDNA product.
KE Type IV collagenase; keloïds; intervertebral disc disease; ds.
KW hypertrophic scars; keloïds; intervertebral disc disease; ds.
OS Homo sapiens.
PN U54923818-A.
PD 08-MAY-1990.
PF 15-MAY-1989; 352069.
PR 15-MAY-1989; US-352069.
PA (UNIW ) UNIV OF WASHINGTON.
PI Goldberg GL, Eisen AZ;
DR WPI; 90-245482/32.
DR R-PSDB; Q05620.
PT Recombinant human type IV collagenase - used in treatment of
PT hypertrophic scars, keloïds and intervertebral disc disease
PS Claim 3; Fig 9; 23pp; English.
CC cDNA clone enables production of type IV collagenase, useful in
CC catalysing cleavage of extracellular matrix macromolecules, and
CC in treatment of hypertrophic scars, keloïds and intervertebral disc
CC disease.
SQ Sequence 660 AA;

```

Query Match      58.6%; Score 51; DB 1; Length 660;  
Best Local Similarity 58.3%; Pred. No. 8.92e-01;  
Matches    7; Conservative    2; Mismatches    2; Indels    1; Gaps    1;

Dbb     239 FNGKE-YNSCTD 249  
      ||| | | : ||:  
QY     62 FNGDEAYDOCTN 73

RESULT    10  
ID R24323 standard; Protein; 73 AA.  
AC R24323;  
DT 18-NOV-1992 (first entry)  
DE Hirudin (HV-3) CRGDWPC.  
KW Restriction sites; thrombin; fibrin; fibrinogen.  
OS Synthetic.  
PN WO9207874-A. PN  
PD 14-MAY-1992.  
PF 23-OCT-1991; G01860.  
PR 24-OCT-1990; GB-023149.  
PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.  
PI Dawson KM, Edwards RM, Fallon A;  
PWI: 92-183627/22.  
DR New proteins comprising active protein and integrin-affinity  
PT sequence - are antithrombotics useful in treating and preventing  
PT myocardial infarction, stroke, pulmonary embolism and deep vein  
PT thrombosis  
PS disclosure, page 43; 10lpp; English.  
CC The sequence given is a synthetic hirudin variant in which the  
CCC amino-acid sequence Cys-Arg-Gly-Asp-Tyr-Pro-Cys has been added to the  
CCC C-terminalus of hirudin HV-3. This CRGDWPC sequence affects the binding  
CCC affinity of the hirudin gene product. Hirudin is an anticoagulating  
CCC agent which binds to thrombin and prevents blood coagulation by  
CCC inhibiting thrombin from catalysing the conversion of fibrinogen to  
CCC fibrin, thus preventing the formation of the protein framework of  
CCC blood clots. There are three principal variants of hirudin (HV-1,  
CCC HV-2 and HV-3). The C-terminal segment of hirudin is essential for  
CCC hirudin-thrombin interaction. Removal of the 5' carboxy-terminal





CC non-ulcerogenic strains.  
SQ Sequence 178 AA;

Query Match 55.2%; Score 48; DB 1; Length 178;  
Best Local Similarity 40.0%; Pred. No. 1.87e+02;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 127 YEGAEYDQC 136  
::: |||  
QY 62 FNGDEAYDQC 71

## RESULT 14

ID R12555 standard; Protein; 241 AA.

AC R12555;

DT 25-SEP-1991 (first entry)

DE Thymidylate phosphohydrolase.

KW Deoxyribonucleic acid; thymidine; dTMP; pyrimidine.

OS Bacillus subtilis CMG356 (ATCC 33234).

PN W09109130-A.

PD 27-JUN-1991.

PF 05-DEC-1990; U06993.

PR 08-DEC-1989; US-448158.

PA (CHEM-) CHEMGEN CORP.

PI Mc Dandliss RL, Anderson DM;

DR WPI; 91-208156/28.

DR N-PSDB; Q12528.

PT Microorganism contg. deoxyribonucleic acid - encoding enzyme

PT causing accumulation of pyrimidine deoxyribonucleoside in

PT recoverable amts

PS Disclosure: Fig 7(a-c); 79pp; English.

CC The protein converts a thymidine deoxyribonucleoside monophosphate to

CC a thymidine deoxyribonucleoside. The DNA sequence encoding this

CC enzyme may be used together with metabolic mutations of heterologous

CC DNA, encoding metabolic enzymes, to engineer cultured cells to

CC express thymidine deoxyribonucleosides (TdNs) in recoverable amts.,

CC providing a fermentation source of TdNs.

CC See also Q12526-27.

CC Sequence 241 AA;

Query Match 55.2%; Score 48; DB 1; Length 241;

Best Local Similarity 40.0%; Pred. No. 1.87e+02;

Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 94 YDNDEFYDC 103

::: |||:::

QY 62 FNGDEAYDQC 71

## RESULT 15

ID R44252 standard; Protein; 411 AA.

AC R44252;

DT 17-JUN-1994 (first entry)

DE Sequence of the envelope protein which belongs to the

DE glycoprotein G set (9G) designated EHV1 gg.

KW Envelope glycoprotein; vaccine; equine herpes virus; immunoassay.

OS equine herpesvirus 1.

PN W09324528-A.

PD 09-DEC-1993.

PF 28-MAY-1993; AU0253.

PR 01-JUN-1992; AU-002716.

PA (UIME) UNIV MELBOURNE.

PI Crabb BS, Studdert MJ;

DR WPI; 93-405733/50.

PT New envelope glyco-proteins of equine herpes virus 1 and 4 -

PT useful in immunoassays to differentiate between these viral

PT types, also vaccines, nucleic acid etc.

PS Claim 10; page 53; 74pp; English.

CC EHV1 gg can be used to capture antibody in immunoassay kits for

CC detecting EHV1, i.e. to detect wild type virus or to establish if a

CC horse has been vaccinated. It is now possible to distinguish

CC between EHV4 (which causes rhinopneumonitis) and EHV1 (which causes

CC abortion). AAs 288-350 is an epitope of EHV1 which is able to elicit

CC a type-specific response. Pref. deletion vaccines for EHV1 lack AAs  
CC 288-350.

SQ Sequence 411 AA;

Query Match 54.0%; Score 47; DB 1; Length 411;

Best Local Similarity 75.0%; Pred. No. 2.38e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 286 FVGDETYD 293

|||::||

QY 62 FNGDEAYD 69

Search completed: Wed May 10 12:34:24 2000

Job time : 8 secs.